# Human Factors Issues in Emergency Response Communication

A dissertation submitted in partial fulfilment of the requirements for the degree of

Master of Science in Applied Psychology
At the
University of Canterbury

By
Annabelle Janiece Munnik
2016

## Contents

## List of Figures

## List of Tables

## Acknowledgements

I want to acknowledge and thank Tait Communications for their support and collaboration with designing the task, and for providing the technology for the experiment. I would also like to sincerely thank them for awarding the Tait Master's Scholarship to me, as it has allowed me to put my time and effort into this research. I would also like to thank my supervisors- Katharina, for your awesome help and support, Deak for your help with coming up with the research and your long distance help, and Graeme for your support and help with collaborating with Tait. Special thanks to Matt Ward for his invaluable knowledge for coding the experiment, and condensing the data down, without him this experiment would not have been possible.

## Abstract

**Objective:** To explore sustained attention in an ecologically valid experiment and to compare two forms of communication technology used by Public Safety Officers.

**Background:** The Sustained Attention to Response Task (SART) is a computer based task designed to measure sustained attention. Participants respond to frequently occurring neutral stimuli and withhold responses to rare target stimuli. Errors of commission (incorrectly responding to target) were traditionally taken as indexes of sustained attention ability. However there is debate in the literature as to whether SART measures sustained attention or ability to inhibit a prepotent motor response (response inhibition theory). A number of hypotheses and research questions were tested, in an ecologically valid setting, to investigate whether SART measures response inhibition or sustained attention, and to test the effects of different types of communication technology on performance.

**Method:** Participants completed a target rich task (high go/low no-go), a target sparse task (low go/high no-go), a verbal recall task, and dual versions of the target rich and target sparse tasks, with the verbal recall task as the secondary task. Participants used a ‘Taser’ to subdue threats (images of people holding guns) on large screens. Participants used either no technology, or one of two forms of radio communication technology used by Public Safety Officers to complete the recall task.

**Results:** Results largely supported the theory that response inhibition is involved in the SART, which is consistent with previous research. There were minimal differences in performance across the technology groups.

**Conclusion:** Results for the traditional computer-based SART have been extended to the present study which employed novel stimuli. Future studies should explore further ways to increase ecological validity of the SART, and investigate whether other perceptual or social processes affect performance when novel stimuli are used.

**Human Factors Issues in Emergency Response Communication**

In every-day life, people have to remain vigilant and sustain their attention. For example, our ancestors had to sustain their attention in the African Savannah to watch for predators. The importance of sustained attention is still prevalent today, such as Public Safety Officers watching for offenders in a busy street filled with pedestrians, or airport security staff watching an X-Ray machine for banned items. Lack of attention is also a critical although frequently overlooked factor in road crashes (Knowles & Tay, 2002).

Traditionally sustained attention has been measured by a well-established low-go/high no-go vigilance task (Helton, Weil, Middlemiss, & Sawers, 2010). Vigilance is the ability to sustain attention during prolonged periods of search and respond to critical and rare occurring stimuli and withhold responses to frequently occurring neutral stimuli (Davies & Parasuraman, 1982; Warm, 1984). Vigilance performance is measured by incorrectly withholding responses to critical or ‘go’ stimuli (error of omission) and incorrectly responding to neutral or ‘no-go’ stimuli (error of commission). Norman Mackworth (1948) first began investigating the area of vigilance after he noticed that well-trained and motivated radar operators during WWII had slower response times and omission errors 30 minutes into their shift. Mackworth (1948) labelled this decrease in performance as time-on-task increases the ‘vigilance decrement’. Typically, participants are far more likely to incorrectly withhold a response than they are to incorrectly respond, and the vigilance decrement is well established in literature (e.g. Grier, Warm, Dember, Matthews, Galinsky, Szalma, & Parasuraman, 2003; Helton & Warm, 2008; Parasuraman, 1979). These traditional low-go/high no-go tasks present target stimuli less than 20% of the time (Helton, Weil, Middlemiss, & Sawers, 2010). Although traditional low-go vigilance paradigms are around 40 minutes in duration (Warm, 1984), the vigilance decrement occurs in harder vigilance tasks (e.g. when stimuli are less salient) within 12 minutes (Temple et al., 2000).

The present study evaluated sustained attention performance in two variations of a modified sustained attention task. While vigilance paradigms are well established in the literature as a measure of sustained attention lapses, an alternative called the Sustained Attention to Response Task (SART), has been suggested (Robertson, Manly, Andrade, Baddeley, & Yiend, 1997). Compared to vigilance tasks, the SART has frequent go stimuli and rare no-go stimuli. However, there is strong evidence to suggest that the SART measures one's ability to inhibit a prepotent motor response, rather than purely sustained attention. The present study aimed to extend previous research investigating what the SART measures by comparing SART performance to performance on an established measure of sustained attention (low-go vigilance tasks). The present study also investigated sustained attention performance in a more ecologically valid setting that reflected constraints of the real-world, and participants were required to interact with radio communication technology.

## The Sustained Attention to Response Task

The Sustained Attention to Response Task (SART) is a modified vigilance task originally designed as a faster method than traditional vigilance tasks to measure sustained attention lapses. The original creators of the SART defined sustained attention as "…the ability to self-sustain mindful, conscious processing of stimuli whose repetitive, non-arousing qualities would otherwise lead to habituation and distraction to other stimuli" (Robertson et al., 1997, p.747). The rationale behind developing the SART was that vigilance tasks can take over ten minutes to reveal lapses in sustained attention, whereas the SART was designed to reveal sustained attention lapses within several minutes. A faster method is useful when measuring sustained attention in clinical populations such as patients with traumatic brain injury (TBI; Robertson et al., 1997). Compared to traditional vigilance tasks, the SART requires participants to respond to frequently occurring neutral stimuli and withhold responses to rare target (no-go) stimuli. Digits one to nine are typically used as stimuli, with each digit randomly displayed 25 times over 4.3 minutes. Font size is varied to ensure participants are processing numbers, not perceptual features. Participants are instructed to withhold responses to digit 3 (target stimuli) and to respond as fast and as accurately as possible to all other digits. Therefore neutral stimuli are presented 89% of the time, and target stimuli 11% (Robertson et al., 1997). Errors of commission, or incorrectly responding to target stimuli, and response time are the metrics used to measure SART performance. Commission errors are interpreted as lapses of sustained attention (Robertson et al., 1997). Healthy participants make commission errors 25% to 50% of the time (Doneva & De Fockert, 2014). A speed-accuracy trade-off (SATO) is a key feature of the SART, in that participants who respond fast also tend to make more commission errors than participants who respond slower (Robertson et al., 1997).

Robertson and colleagues (1997) argued that the SART measures sustained attention. Participants in the 1997 study were healthy controls and patients with traumatic brain injury, who completed the SART and a Cognitive Failures Questionnaire (CFQ). The authors sampled items in the CFQ that they believed related to attention, memory and action slips. Therefore arguably some of these items may not be directly related to lapses in sustained attention (Smilek, Carriere & Cheyne, 2010). However, Robertson and colleagues (1997) found a positive correlation between SART performance and the CFQ ($r$ = 0.27), which was interpreted as participants that performed worse on the SART were more likely to have everyday attention, memory and action slips. The authors therefore argued that the SART measures lapses in sustained attention. Although there has been debate over the validity of correlating the SART and the CFQ, SART has also been correlated with more specific measures of sustained attention using traumatic brain injury patients, such as the Attention-

Related Cognitive Errors Scale (*r* = .23; Smilek, Carrier, & Cheyne, 2010). The SART is used in a wide range of clinical populations as a measure of sustained attention lapses, such as patients with depression (Farrin, Hull, Unwin, Wykes, & David, 2003), children with attention deficit/hyperactivity disorder (ADHD; Bellgrove, Hawi, Kirley, Gill, & Robertson, 2005), and narcolepsy patients (Fronczek, Middelkoop, Van Dijk, & Lammers, 2006). The SART is also used to determine whether interventions to improve sustained attention are effective (Seli, Jonker, Solman, Cheyne, & Smilek, 2013). The SART was designed to measure sustained attention and has therefore been used for this purpose in past research.

**Mindlessness theory**

Proponents of mindlessness theory of sustained attention argue that lapses in sustained attention on the SART are caused by mind wandering (Robertson et al., 1997). Participants must maintain endogenous attention (voluntarily directed attention) to respond correctly in the SART. The objective monotony of the SART causes a decrease in endogenous attention and the mind wanders, therefore SART performance declines (Robertson et al., 1997). For example, participants make fewer errors of commission as the proportion of target (no-go) stimuli increases (Manly, Robertson, Galloway, & Hawkins, 1999). Mindlessness theory argues that the increased presentation of target stimuli (50% compared to 11%) provides external support for attention. In other words, more target stimuli externally captures participants attention, rather than relying on participants to voluntarily direct their attention. External support for attention therefore redirects attention to the task, leading to improved SART performance (Manly, Robertson, Galloway, & Hawkins, 1999). Therefore, providing exogenous support for attention, even when unrelated to the SART, should increase SART performance (Manly et al., 2004).

A modified version of the SART, known as $SART_{fixed}$, has been used to test the theory that SART performance is caused by mind wandering. The $SART_{fixed}$ is a modified version of the original SART in which the sequence of digits is predictable, and therefore one can predict when the target stimuli will appear. Healthy controls and traumatic brain injury (TBI) patients completed the $SART_{fixed}$. TBI patients did not slow their responses to stimuli preceding target stimuli in the $SART_{fixed}$, whereas healthy controls slowed their responses in anticipation for the target stimuli they had to withhold a response to. TBI patients made more commission errors than healthy controls. The authors argued that TBI patients had damage to the area of the brain which maintains endogenous control required for sustaining attention, whereas healthy controls did not. Therefore TBI patients $SART_{fixed}$ performance reflects lapses in sustained attention (Dockree, Kelly, Roche, Hogan, Reilly, & Robertson, 2004). However, the authors admit that the original SART may confound sustained attention and inhibitory control (Dockree et al., 2004). Therefore the original SART with random presentation of digits may reflect inability to inhibit a response, and this is mitigated in the $SART_{fixed}$ by having a predictable sequence of digits.

Measuring task-unrelated thoughts is a method used to test the mindlessness theory of sustained attention. Task unrelated thought are mind wandering or ‘zoning out’. In a modified SART which used grey and white squares as stimuli, participants made most of their errors in blocks in which they reported high levels of task-unrelated thoughts (as measured by asking participants what they were thinking about) (Smallwood et al., 2004). Contrary to mindlessness theory, many studies have shown participants report task-related thoughts despite committing commission errors, which suggests that mind wandering may not be the cause of commission errors in the SART (Carter, Russell, & Helton, 2013; Head & Helton, 2012; Helton, Kern, & Walker, 2009).

**Does the SART measure sustained attention?**

Contrary to the original authors of the SART (Manly et al., 1997), other researchers suggest that the high-go/low no-go SART measures one’s ability to inhibit a prepotent motor response, rather than purely measuring lapses in sustained attention (e.g. Carter, Russell, & Helton, 2013; Head & Helton, 2014; Helton, 2009; Helton, Weil, Middlemiss, & Sawers, 2010). Although the authors of the SART have acknowledged that inhibitory processes are involved in the SART (e.g. Dockree et al., 2004; Dockree et al., 2006), they continue to focus on perceptual processes (such as using cues to provide support for exogenous attention) rather than motor control or strategy concerns (Helton, Weil, Middlemiss, & Sawers, 2010). This debate has important implications, as the proponents of mindlessness theory assume SART measures sustained attention when there is evidence to suggest it does not (e.g. Helton, Weil, Middlemiss, & Sawers, 2010; Helton & Russell, 2011) and also interventions to improve sustained attention use SART to track intervention efficacy (Carter, Russell, & Helton, 2013). If the SART does not measure sustained attention well, then it should not be used to test theories of sustained attention or to track sustained attention intervention efficacy. The SART may measure one’s ability to inhibit a prepotent motor response (Head & Helton, 2014).

The response inhibition theory proposes that participants respond to go stimuli repeatedly and in short succession, and they therefore develop a prepotent motor response. Participants develop a motor response of pushing the spacebar/mouse due to the high frequency of go trials, and this becomes a dominant response that is difficult to inhibit even if a response is not required. Therefore when rare no-go stimuli appear, participants commit a commission error as they cannot inhibit their prepotent motor response, as compared to a commission errors indexing lapses of sustained attention (Head & Helton, 2012, 2013). For example, participants are aware of stimuli but report being unable to inhibit their response (Cheyne, Carriere, & Smilek, 2009).

The prepotent motor response may develop as participants struggle to fulfil task requirements of responding as fast and as accurately as possible (Head & Helton, 2014). Due to this, participants choose to implement one of two strategies depending on the utility of each strategy. They can either respond fast to all stimuli, which is less accurate, or respond accurately to all stimuli which slows response time. Due to the nature of the SART and the fact that there are few target stimuli that require withheld responses, strategically responding fast to every stimulus is overall mostly beneficial for performance, and will become a prepotent motor response. Therefore high-go/low no-go tasks bias participants towards responding fast in sacrifice for accuracy, which is known as the speed-accuracy trade-off (Helton, Weil, Middlemiss, & Sawers, 2010). However this strategy will result in commission errors when rare target stimuli occur. When a commission error is made, participants strategically slow responses to enable checking of stimuli before responding (Peebles & Bothell, 2004), and may cycle back to the fast responding strategy when the utility of it is maximized. Although Manly and colleagues (2000) are proponents of mindlessness theory, they found that reaction times before a commission error were faster than response times before a correct withhold, and that reaction times were slower after a commission error, compared to a correct withhold. Peebles and Bothell (2004) argue that this result reflects the response strategy theory, as once a commission error is made the response strategy is switched to slow down speed of responses to enable checking of stimuli.

Support for the response inhibition theory of SART has been derived from paradigms comparing SART performance to performance on traditional vigilance tasks. Traditionally formatted tasks (TFT) are low-go/high no-go vigilance tasks in which participants respond to rare occurring stimuli and withhold responses to frequent occurring neutral stimuli. TFT are a well-documented measure of sustained attention (e.g. Helton & Warm, 2008; Parasuraman, 1979). TFT (low-go tasks) do not require participants to repeatedly respond, therefore they do not provide the opportunity for a prepotent motor response to develop. However SART is meant to bias participants towards sacrificing accuracy for speed, known as the speed-accuracy trade-off (Helton, Weil, Middlemiss, & Sawers, 2010). Comparing performance on perceptually equivalent SART and TFT, in which the same stimuli are used (but percentage of target stimuli differs) can partial out the prepotent motor response aspect of SART performance (Carter, Russell, & Helton, 2013). According to mindlessness theory, any differences in SART and TFT performance are attributed to the change in external perceptions by changing response format (even if stimuli are the same). It is not very plausible that changing response format would lead to differences in performance if SART is truly a measure of sustained attention (Carter, Russell, & Helton, 2013).

Previous studies comparing the SART with traditional vigilance tasks have concluded that the SART somewhat measures response inhibition. More commission errors are made in a SART than a perceptually equivalent TFT (Carter, Russell, & Helton, 2013; Dippel, Chmielewski, Muckschel, & Beste, 2015; Helton, Weil, Middlemiss, & Sawers, 2010; Helton & Russell, 2011). It was argued that the SART biases participants towards responding incorrectly to target stimuli, whereas the TFT does not, and that this is likely due to inability to inhibit a prepotent motor response. This interpretation is further supported by the results that a speed-accuracy trade-off occurs in SART’s but not in perceptually equivalent TFT’s (Helton, Weil, Middlemiss, & Sawers, 2010; Stevenson, Russell, & Helton 2011). Response inhibition authors argue that this is because SART biases participants towards responding fast at the sacrifice of accuracy, whereas TFT does not. This is further supported by the fact that response times were faster in the SART than perceptually equivalent TFT (Dippel, Chmielewski, Muckschel, & Beste, 2015; Helton, & Russell, 2011; Helton, Weil, Middlemiss, & Sawers, 2010; Stevenson, Russell, & Helton, 2011). The majority of studies have found no difference in omission errors between perceptually equivalent SART and TFT (Carter, Russell, Helton, 2013; Dippel, Chmielewski, Muckschel, & Beste, 2015; Helton & Russell, 2011), although one study found more omission errors in SART than TFT (Helton, Weil, Middlemiss, & Sawers, 2010). Researchers argue that omission errors are a better measure of lapses in sustained attention in the SART, as omission errors are an accepted metric for lapses in sustained attention in TFT (Carter, Russell, & Helton, 2013). However, the above results provide strong support for the argument that the SART confounds response inhibition with sustained attention.

Reliable warning cues improve SART performance, which supports the response inhibition theory of SART. One could argue from a mindlessness perspective that including cues captures exogenous attention and reorients attention back to the task. However, it is not the mere fact of including cues which improves performance as participants who have unreliable cues, which do not correctly predict stimuli, make more commission errors than participants who have cues which reliably predict stimuli (Finkbeiner, Wilson, Russell, & Helton, 2015; Helton, Head, & Russell, 2011). There is also a speed-accuracy trade-off with unreliable cues, but not with reliable cues (Helton, Head, & Russell, 2011). Therefore reliable cues allow participants to better cycle between their response strategies and control their prepotent motor response, which reduces the speed-accuracy trade-off and improves performance. Participants who have reliable cues make more omission errors than participants with unreliable cues or no cues (Helton, Head, & Russell, 2011). Omission errors could be tactical rest-stops for participants to monitor and control their prepotent response (Helton & Russell, 2011). This perspective is supported by results which show participants with reliable warning cues had slower reaction times to stimuli preceding omission errors (compared to correct response), whereas participants with unreliable cues or no cues had faster reaction time preceding omission errors (Helton, Head, & Russell, 2011). Overall, studies using warning cues support the response inhibition perspective of SART.

Forcing participants to slow their responses can reduce the speed-accuracy trade-off and increase SART performance, which supports the response inhibition theory of SART performance. Manual selection requires participants to manually select stimuli, as compared to automatic selection of stimuli where the participant merely pushes a button to respond. When participants are required to manually select stimuli with a joystick and stimuli location on a computer screen is unpredictable, there is a non-significant correlation between errors of commission and response time (Head & Helton, 2014). The speed-accuracy trade-off only disappears when both manual selection and unpredictable location are used, as compared to using either manual selection or unpredictable location (Head & Helton, 2013). Manual selection also leads to fewer commission errors than automatic selection (Head & Helton, 2013). Manual selection and unpredictable stimuli location essentially forces participants to slow their responses which reduces the prepotent motor response, so although errors of commission are still committed, performance increases and the speed-accuracy trade-off is reduced (Head & Helton, 2014). Response time can also be manipulated by requiring participants to respond in sync with a metronome. Participants that were delayed to 800ms after stimuli presentation, due to the metronome, had fewer errors of omission and commission than participants responding 400ms or 600ms after stimuli presentation, suggesting there may be a threshold for response time in which performance increases (Seli, Jonker, Solman, Cheyne, & Smilek, 2013). Response time can also be altered by emphasising accuracy when giving instructions rather than speed (Seli, Cheyne, & Smilek, 2012). These results indicate that manipulating or delaying response time decreases the speed-accuracy trade-off and increases SART performance, and this is independent of sustained attention ability.

Manual selection and unpredictable stimuli location can lead to more errors of omission than automatic selection and predictable location (Head & Helton, 2013; Helton, Weil, Middlemiss, & Sawers, 2010). Making location unpredictable essentially makes a task harder by increasing cognitive load, and therefore participants may be more hesitant to respond (Helton, Weil, Middlemiss, & Sawers, 2010) and task requirements of scanning the screen to locate stimuli can lead to failures to respond (Head & Helton, 2013). This conclusion is supported by the result of a positive correlation between response time and omission errors when participants were required to manually select stimuli and location was unpredictable. This suggests that participants found it hard to manually select stimuli while also searching for stimuli, and that they ran out of time to respond (Head & Helton, 2014).

**Moving towards ecologically valid SART paradigms**

Literature has recently begun to focus on more ecologically valid high-go SART paradigms. Ecological validity can be defined in terms of the SART reflecting sustained attention lapses in patient populations who are administered the SART, such as ADHD and traumatic brain injury patients (Smilek, Carriere, & Cheyne, 2010). Another method of approaching ecological validity is to consider whether stimuli and task demands reflect the nature of the real-world. The traditional computer based SART with number stimuli does not reflect real-life SART constraints, such as attending to dynamic and novel stimuli in a changing environment (Head & Helton, 2015). This is critical to understand, particularly as the SART is not only used to diagnose patients with deficiencies in sustained attention, but also to test theories of sustained attention for normal populations and to measure the efficacy of sustained attention interventions (Carter, Russell, & Helton, 2013). Although the current computer based SARTs do reflect some real-world constraints, such as those faced by radar operators investigated by Mackworth (1948), and the effects of manual selection and unpredictable location (Head & Helton, 2014), this is still somewhat limited in terms of ecological validity. The tasks employed to measure sustained attention should reflect constraints of the real world to ensure accurate conclusions are made. For example, Police Officers use radio communication technology for the duration of their shift while also attending to critical target stimuli in their environment, such as watching out for a specific number plate of an offending vehicle. How would this dual-task scenario with multiple demands on attention affect their ability to sustain their attention? To answer such questions it is also important to consider whether response inhibition is part of performance in more ecologically valid SART paradigms.

**Pictures as stimuli.** Several experiments have used pictures as stimuli, which is arguably more ecologically valid than number stimuli. Using pictures of cities and nature as stimuli led to more errors of commission made by healthy participants than commission errors made by brain injured patients in the traditional SART (Head & Helton, 2012; Robertson et al., 1997). Rather than providing exogenous support for attention and decreasing errors as mindlessness theory would predict, natural scene stimuli were interpreted as placing more cognitive load on participants thereby disrupting their ability to inhibit response (Head & Helton, 2012). Pictures of spiders and neutral objects (such as chairs) have been used as go and no-go stimuli. Spider pictures were designed to induce task-relevant anxiety, which is hypothesized to improve ability to inhibit a prepotent response (Wilson, Russell, & Helton, 2015). There were fewer errors of commission in the picture SART compared to the traditional number SART. It is unclear whether a decrease in commission errors was caused by the anxiety spiders induced or because spider stimuli were arguably more salient than the traditional number stimuli (Wilson, Russell, & Helton, 2015). Therefore, it is not entirely clear how pictures influence SART performance. It appears to be a mixture of how salient the stimuli are, if stimuli induce anxiety, and how much cognitive load stimuli induce.

**Shooter scenario.** SART performance has been applied to the real world example of friendly fire incidents. Wilson and colleagues (2015) investigated if high-go/low no-go SART performance reflects ability to inhibit a prepotent motor response in a more realistic situation in which participants had to physically shoot moving research assistants with a near infrared emitter gun. In experiment one, proportion of target stimuli was manipulated. Participants were instructed to search rooms in a building and respond to foes (wearing hats), and withhold responses to friendly people. Participants completed this task three times, and proportion of foes was either 89%, 50% or 11%. Response time was not recorded. Participants were more likely to make errors of commission when the proportion of foes were higher. The authors interpreted this as when participants were required to respond more frequently, they developed a prepotent motor response which was difficult to inhibit. Subjective task focus reports were higher as percentage of foes increased, which indicates participants were focused on the task even though they were committing commission errors. These results suggest response inhibition is more likely than mindlessness to explain SART performance (Wilson et al., 2015).

In experiment two, participants completed a modified version of experiment one while also completing computer based versions of sustained attention tasks. Participants completed the firearm high-go/low no-go task (target rich equivalent to the SART), firearm low-go/high no-go task (target sparse equivalent to traditional vigilance tasks), and the computer equivalents of each of these tasks. In the firearm tasks, participants leaned on a structure, and the emitter gun was aimed at the area in which research assistants would appear (location predictable and gun aimed at everyone), and foes were indicated by the direction a balaclava was worn. The research assistant appeared in the designated area every 1.5 seconds. The computer tasks used numbers as stimuli, and participants responded with trigger press on gun. Subjective workload was assessed through the NASA-Task Load Index (NASA-TLX;

Hart & Staveland, 1988). More commission errors were committed in the target rich tasks than target sparse tasks. A speed-accuracy trade-off was evident in both the computer and firearm target rich tasks. More omission errors were made in the firearm target rich task than the comparative computer target rich task. The authors suggested that participants use omission errors as strategic ‘pauses’ due to the high cognitive and physical demands of the emitter gun SART. There was no difference in omission errors between the target rich and target sparse tasks. Subjective workload ratings were higher in the target rich tasks than the target sparse tasks, suggesting that the target rich tasks require more cognitive resources than a vigilance task. The authors argued that this is because target rich tasks require response inhibition whereas target sparse tasks do not. This more realistic emitter gun scenario therefore provides support for the response inhibition theory, indicating that friendly fire incidents may be caused by inability to withhold a prepotent motor response (Wilson, Head, de Joux, Finkbeiner, & Helton, 2015).

**Dual-task performance.** Several studies have investigated SART performance in a dual-task scenario, thereby attempting to assess SART performance in more realistic scenarios. In reality, people often have to do more than one thing at a time. For example, someone may need to sustain their attention when looking for dangers on the road whilst simultaneously making a phone call on their Bluetooth set. A Police Officer may need to radio dispatch whilst looking for an offender in a red hat while driving around a certain location. The multiple resource theory of dual-task performance states that performance decreases when two tasks share cognitive resources (Wickens, 2002, 2008). Proponents of the response inhibition theory of SART argue that adding a secondary task will decrease SART performance if the secondary task has a cognitive load due to competition for limited resources (Head & Helton, 2014). However the mindlessness theory of SART suggests that adding a secondary task will capture attention (or provide exogenous support for attention) thereby making SART less monotonous, leading to an increase in performance (Manly et al., 2004) and high-task load (adding a secondary task) should decrease the likelihood of mind wandering which leads to improved SART performance (Smallwood & Schooler, 2006).

Previous studies assessing dual-task performance in SART paradigms support the response inhibition account of SART. When the secondary task completed alongside a SART is difficult or requires multiple processing stages, such as encoding and maintaining words (Head & Helton, 2014) or maintaining and manipulating numbers (Grandjean & Collette, 2011), as compared to simply maintaining numbers in memory (Doneva & De Fockert, 2014), errors of commission increase. This suggests that when more resources are required to complete the secondary task, inhibiting a prepotent response becomes more difficult, which supports the response inhibition theory of SART (Head & Helton, 2014). Errors of omission also increase, indicating that participants take strategic breaks during perceptually difficult tasks due to limited cognitive resources (Doneva & De Fockert, 2014). However, the above studies used a computer SART with numbers or coloured figures as stimuli, and the secondary tasks were relatively simple which does not necessarily reflect real-world constraints.

**Aim of present research**

There is currently a gap in the literature looking at SART or high-go/low no-go task performance in ecologically valid settings. The importance of understanding SART performance in the real world can be illustrated with an example of a Police Officer. The officer may be situated outside a scene, tasked with keeping watch for a group of gang members. The street may be cordoned off so there will be few neutral passer-bys. The offenders are the stimuli in which police must respond to, whilst they must inhibit responses to passer-bys. While the Officer is doing this, they will be receiving instructions, updates and welfare checks from dispatch over the radio. Understanding SART performance in realistic situations in which there are multiple demands on attention is critical as researchers need to understand the causes of SART performance, and also sustained attention in general, to ensure appropriate steps are taken to mitigate the negative effects.

The current research aims to investigate SART performance in a more ecologically valid setting in which precise performance metrics can be obtained, and to compare SART performance to a more traditionally accepted measure of sustained attention (low-go vigilance task). The SART in the current study will be referred to as the ‘target rich task’, and the vigilance, or low-go/high no-go task, as the ‘target sparse task’. To investigate this participants stood in the middle of three large screens, and used a Taser like joystick to respond to foe stimuli, who were realistic drawings of people holding guns. Participants were instructed to withhold responses to friendly stimuli, who were realistic drawings of people not holding guns. The percentage of foe targets was 89% for the target rich task, and 11% for the target sparse task. The stimuli used for the target rich and target sparse tasks were therefore perceptually equivalent, and the target sparse was included as a task as low-go/high no-go tasks are an accepted measure of sustained attention (Helton & Warm, 2008), and therefore allowed for comparison to performance on the target rich task. Participants also completed a modified version of the NASA-TLX (Hart & Staveland, 1988) after each task to assess subjective workload.

Another objective of this current research was to compare two forms of communication technology used by Public Safety agencies developed by Tait Communications. Including these two technologies in the study arguably increased the ecologically validity of the secondary task completed alongside the target rich and target sparse tasks. The traditional Radio required pushing a button to radio dispatch. The ‘Mobility App’ is an application for a mobile phone which the user must log on to before radioing. The user can choose which group they want to radio, and then hold a ‘push to talk’ button for the duration of their radio message to communicate with the selected group. The verbal recall task in the current study involved participants radioing an associated word back (using either the Radio, Mobility App, or no technology) when they heard a word through the headphones, and they recalled the words they remembered at the end of the task. Participants completed the target rich single task, target sparse single task, the verbal recall task, target sparse dual task (verbal recall and target rich task), and target sparse dual task (verbal recall and target sparse task). It was deemed most appropriate to test the radio technology in the above Taser and verbal recall scenario, rather than in a paperwork or office situation, as the importance of having effective technology that minimally interferes with cognitive processes is more so in tasks where officer safety is a concern, such as the above Taser scenario.

## Hypotheses and research questions

It is expected that the more ecologically valid target rich (high-go/low no-go) tasks will extend results found in past studies using the SART high-go paradigm that indicates response inhibition plays a role in performance.

## Response time

***Research question 1a***. Will a speed-accuracy trade-off be evident in the target rich tasks? Making location stimuli unpredictable and requiring participants to manually select stimuli reduces the speed-accuracy trade-off (Helton & Head, 2014). The current study requires participants to manually select stimuli, and stimuli location is unpredictable. However, a speed-accuracy trade-off (as evidenced by a negative correlation between commission errors and response time) was found in the more realistic emitter gun SART, although stimuli location was not unpredictable in this study (Wilson et al., 2015). Therefore a speed-accuracy trade-off may not be found in the current target rich tasks if response time is delayed enough for participants to control their prepotent motor response.

***Research question 1b***. Is there a relationship between response time and errors of omission in the target rich tasks? Previous studies have found that requiring participants to manually select stimuli when location of stimuli is unpredictable leads to a positive correlation between errors of omission and response time, and it was suggested that participants found manual selection hard and missed responding to stimuli due to the additional cognitive load of having to look for where stimuli were (Head & Helton, 2013, 2014). It is possible that participants may find it hard to respond in time in the present study, but it is unclear whether there will be a relationship between errors of omission and response time, as participants may not find the present more realistic target rich tasks difficult.

***Hypothesis 1a.*** There should be no speed-accuracy trade-off in the target sparse tasks, as low-go/high no-go tasks do not bias participants towards fast responding (Helton, Weil, Middlemiss, & Sawers, 2010; Stevenson, Russell, & Helton 2011).

***Hypothesis 1b.*** Response time will be faster in target rich tasks compared to target sparse tasks, as high-go/low no-go tasks have been shown to bias participants towards fast responding (Dippel, Chmielewski, Muckschel, & Beste, 2015; Helton, & Russell, 2011; Helton, Weil, Middlemiss, & Sawers, 2010; Stevenson, Russell, & Helton, 2011).

**Errors of commission and omission.**

***Research question 2.*** What is the difference in omission errors between the target rich tasks and target sparse tasks? Most studies have found no difference in omission errors between perceptually equivalent high-go tasks and low-go tasks (Carter, Russell, Helton, 2013; Dippel, Chmielewski, Muckschel, & Beste, 2015; Helton & Russell, 2011), although high-go tasks have been shown to have more omission errors than perceptually equivalent low-go tasks (Helton, Weil, Middlemiss, & Sawers, 2010). The response inhibition perspective of high-go SART performance focuses on commission errors, therefore it has been suggested omission errors are more likely to reflect actual lapses in sustained attention (Helton, Weil, Middlemiss, & Sawers, 2010).

***Hypothesis 2a.*** There will be more errors of commission in the target rich tasks than target sparse tasks (Helton & Russell, 2011). Previous studies have found this result, and argued that target rich SART tasks introduce response inhibition demands (therefore more likely to incorrectly respond to no-go stimuli) whereas the traditional target sparse vigilance tasks do not (Carter, Russell, & Helton, 2013; Dippel, Chmielewski, Muckschel, & Beste, 2015; Helton, Weil, Middlemiss, & Sawers, 2010; Helton & Russell, 2011).

***Hypothesis 2b.*** More errors of commission will be committed in the dual tasks than single tasks. The dual tasks place additional cognitive load on participants, as they have to also complete a verbal task alongside the sustained attention task, and therefore performance should get worse on the dual tasks. In terms of the target rich tasks, difficult secondary tasks have been shown to disrupt one's ability to inhibit a prepotent response, which is evidenced by an increase in commission errors on dual task compared to single (Grandjean & Collette, 2011;Head & Helton, 2014 ). This suggests that dual-task scenario places additional demands on limited cognitive resources, according to response inhibition theory.

***Hypothesis 2c.*** More errors of omission will be committed in the dual-task than single task, as the dual tasks place additional cognitive load on participants, which should increase omission errors.

***Hypothesis 2d.*** The people using the Mobility App will commit more commission errors than the people using the Radio, as the Mobility App is hypothesized to place more cognitive load on participants as it involves checking the interface and ensuring the radio connection remains established. This additional cognitive load is therefore hypothesized to disrupt participant's ability to inhibit their prepotent motor response, as evidenced by more commission errors.

***Hypothesis 2e.*** The people using the Mobility App will commit more errors of omission than the people using the Radio. This is hypothesized for the same reasons as hypothesis 2d. As the Mobility App requires participants to avert their eyes from the sustained attention task to check the radio connection is still established, more errors of omission will be committed than those people using the Radio.

***Hypothesis 2f.*** Participants in the control condition will commit the fewest errors of commission and omission, as they are not required to interact with any technology that could disrupt their performance.

**Word recall**

***Hypothesis 3a.*** More words will be recalled in the verbal recall task than the dual sustained attention tasks.

***Hypothesis 3b*****.** People using the Mobility App will recall fewer words than those using the Radio. This is for the same reasons as outlined above, as the Mobility App is hypothesized to place more cognitive load on participants (as it involves checking the interface and ensuring the radio connection remains established), and will therefore interfere with participants performance.

***Hypothesis 3c.*** Those using no technology in the control condition will recall more words than participants required to use technology (the Mobility App or Radio). This is because participants using technology have additional cognitive load added to the verbal recall task (they must not only recall words, but also interact with technology). Essentially the control condition should be easiest, which should lead to more words recalled than those using technology.

**Workload ratings.**

***Hypothesis 4a***. Workload ratings will be higher for the target rich tasks than target sparse tasks. This result was found in the realistic firearm SART (Wilson et al., 2015), and the authors suggested that the high-go SART requires more cognitive resources than a low-go vigilance task as SART also requires participants to inhibit a prepotent motor response.

***Hypothesis 4b*.** Workload ratings will be higher in the dual tasks than single tasks, as the dual tasks are hypothesized to add additional cognitive load to the single tasks, and will therefore have higher subjective workload ratings.

***Hypothesis 4c***. Workload ratings will be higher for people using the Mobility App than people using the Radio, as the App is hypothesized to require more cognitive effort to correctly use (checking the connection for example) than simply pressing the button on the Radio set.

***Hypothesis 4d.*** The control condition will have the lowest workload ratings, as it is essentially the easier condition, as the technology conditions require interactions with technology whereas the control condition only requires participants to say words aloud.

# Method

## Participants

Sixty two University of Canterbury Undergraduate Psychology students were recruited as participants, and received course credit in return for their participation (and one fourth year engineering student who received a $10 voucher for participation). This was deemed acceptable as SART performance is relatively stable across age, gender and education (Chan, 2001). Participants were fluent in English, and had normal (or corrected-to-normal) hearing and vision. There were 10 males and 53 females, and their ages ranged from 18 to 45 ($M$ = 20.25, $SD$ = 4.19). Data for three participants was removed due to an interruption during the experiment, or the technology did not record responses. Overall there were twenty participants in each technology group. The Mobility App group had 4 males and 16 females, Radio group had 3 males and 17 females, and control group had 3 males and 17 females.

## Materials

**Hardware.** The experiment took place in the Vision Space in the Human Interface Technology Lab NZ, at the University of Canterbury. This space has a middle screen, with two screens at 60 degree angles to the middle screen. Each screen is 244cm wide, 180cm high and raised 60cm above the ground. See Figure 1 for the room layout. Participants were instructed to stand in the middle of the screens on a white cross on the floor that was 160cm from the centre screen. Three NEC LT265 projectors projected images onto the screens. A desk was placed 3.5 metres away from the centre cross, and participants had to return to this desk after each task to complete the workload questionnaire and/or recall words.

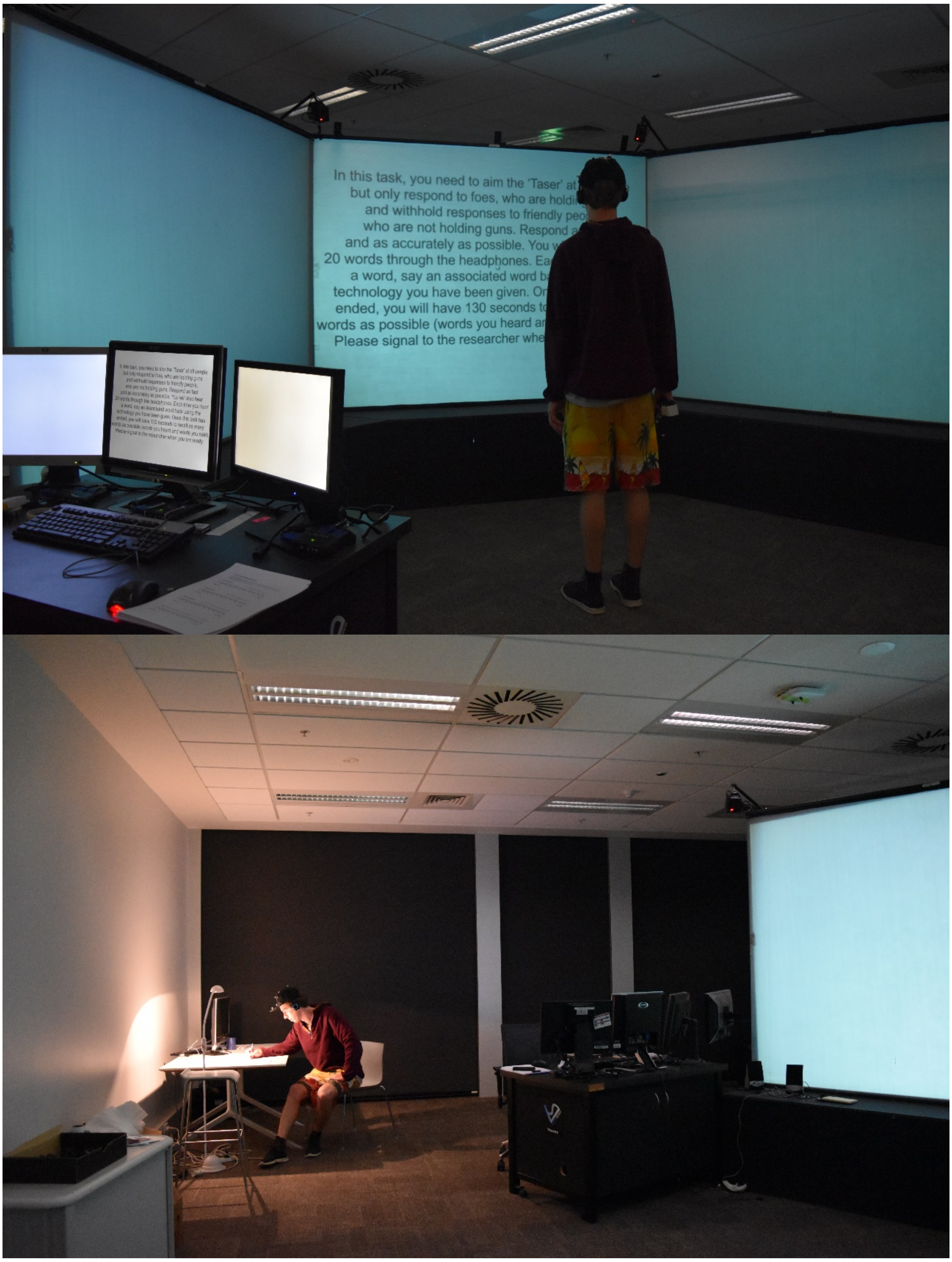


*Figure 1.* The Vision Space

Participants wore a black cap (with markers attached) over the headphones, and were under the impression this recorded their head movements. Participants used a Flystick (Flying Joystick; Abasa, Didier, Tazi, & Mallem, 2007) to respond by pulling a trigger at the front of the flystick. This flystick looks similar to Taser guns used by Public Safety agencies around the world, and the flystick will be referred to as the ‘Taser’. The ‘Taser’ also had markers attached for tracking movement, although tracking was not used in the current experiment. The hat and ‘Taser’ can be seen in Figure 2.

*Figure 2.* The ‘Taser’ and hat

**Radio communication technology.** The Mobility Application (Mobility App) was on an iPhone 5 in a black case. Participants had to swipe into the iPhone, and tap on the app icon once at the beginning of the experiment to get into the Mobility App (this was done before testing began). There was only one group that they could radio, which was labelled “UC2”. Participants had to tap on “UC2” to establish a connection, and hold down the ‘push to talk’ button at the bottom of the screen for the duration of their message to radio a word through. If the participant did not radio a word through within 30 seconds after establishing a connection with “UC2”, it would time out, and they would have to tap on “UC2” again to establish the connection. The researcher used s Samsung Galaxy S3 Mini to listen to words radioed by the participant.

The Radio was comprised of two parts, as seen in Figure 3. The main part has a screen, volume control, and options to change radio channels. The radio channel was pre-set to “A1”, which radioed the participants’ message directly to the researcher. Participants were not required to interact with this main part of the Radio. The second part of the Radio is a microphone that is clipped to the user’s lapel. There was a button approximately 5cm long on the side of the microphone, and participants had to hold this down for the duration of their radio message to send a message through to the researcher. The researcher also had a Radio which they plugged earphones in for the duration of the experiment to listen for the words radioed through.


*Figure 3.* The Radio

**Stimuli.** Two hundred and twenty-five realistic images of people were created using Adobe Fuse. There were 113 male stimuli created, and 112 female stimuli. Six different body types and ethnicities were used between 18-19 times for each gender of stimuli. Height, weight and facial features were randomised for each stimulus. A wide variety of clothes, hairstyle and shoes were chosen by the researcher, to ensure stimuli encompassed a wide range of people. Adobe Photoshop was used to give stimuli the ‘arms down’ pose, and the facial expression of ‘indifferent’. Background shadows were removed. For the target rich task, 89% of these stimuli randomly had a 'colt' gun placed in their hand. For the target sparse task, a random 11% of these stimuli had a 'colt' gun placed in their hand. Guns were randomly assigned to the right or left hand, with an equal number across foe stimuli. This same pool of 225 images of people were used for all target rich and target sparse tasks, the only difference being the proportion of people which had guns. Stimuli were also created for target rich and target sparse practice trials. These practice trials consisted of 18 stimuli which were not used in the actual tasks. In the target rich practice trial, two stimuli were friendly, and in the target sparse trial, 16 stimuli were friendly. Example of stimuli can be seen in Figure 4. Unity software was used to build the experiment (Helgason, Francis, & Ante, 2005).

*Figure 4.* Example of friendly and foe stimuli used

**Audio tracks.** Three different word lists, containing 20 words each, were used to ensure practice effects for words did not confound results. Word lists were taken from Head and Helton (2014), who generated the word lists from a pool of 925 nouns from Paivio and colleagues (1968). Generated words were randomly assigned to one of the three lists. To ensure words were similar and equally memorable, parameters for word selection were: 2 syllables and 5-7 letters long, Kucera-Francis word frequency of 0-30, concreteness rating of 6-7, imagery rating of 5-7 and meaningfulness rating of 6-8 (Head & Helton, 2014). Refer to Appendix A for the word lists. Words was recorded on a Samsung Galaxy S6 by a female native New Zealand speaker. Camtasia studio was used to create the audio tracks. To create the two scrambled word lists, each word was spliced into three parts using WavePad audio editing software. This led to 180 spliced audio tracks. One scrambled word was created by randomly combining three spliced tracks. Words were played to participants through a Logitech H600 Wireless headset at volume 65 on the desktop computer. The audio tracks were the same length as the target rich and target sparse tasks (7 minutes and 37 seconds long). There was 37 seconds silence at the beginning and end of each audio track. A word was played every 20 seconds. A practice audio track was also created using four words from Paivio and colleagues (1968) that were not used in the actual word lists. The practice audio track had fifteen seconds of silence at the beginning and end, and a word said every 20 seconds. A Samsung Galaxy S6 was used to record the words participants radioed or said back.

**Questionnaire.** A modified NASA-TLX was employed (Hart & Staveland, 1988) as a measure of subjective workload. Participants were asked to rate four dimensions of the original six dimensions of the NASA-TLX on a visual analog scale from 0-100: mental demand (how mentally demanding was the task?), physical demand (how physically demanding was the task), temporal demand (how hurried or rushed was the pace of the task?)

and effort (how hard did you have to work to accomplish your level of performance?). Higher scores reflect higher subjective workload. Participants were instructed to think of the task they had just completed when completing the questionnaire and to mark a cross which best reflects their answer. The centre of the cross was taken as participants score. Please refer to Appendix B for a copy of the NASA-TLX used.

## Design

**Independent variables.** This study used a 2 X 2 X 3 mixed designed, and there were three independent variables. Task type was a within subjects independent variable with two levels; target rich (89% foes, 11% friendly) and target sparse (11% foes, 89% friendly). The second independent variable was cognitive load and was a within subjects variable with two levels; single load and dual load. The secondary task added to the single tasks was a verbal recall task. These two independent variables led to four tasks participants had to complete; target rich single task, target rich dual task, target sparse single task, and target sparse dual task. Participants also completed the verbal recall task by itself, to get a base level of performance for word recall. The third independent variable was the type of technology used for the verbal recall task. This was a between subjects variable with three levels; control (say words aloud), the Mobility App and the Radio. Twenty participants were randomly assigned to each technology group. This was a between subjects factor to reduce the potential confound of practice effects, as otherwise participants would have to complete the sustained attention tasks many times, which has been shown to improve performance (Whyte, Grieb-Neff, Gantz, & Polansky, 2006).

**Dependent variables.** Dependent variables were response time, errors of commission, errors of omission, number of words recalled, number of associations made in the verbal recall tasks, and subjective workload ratings. Errors of omission are incorrectly withholding a response to a foe, and errors of commission are incorrectly responding to a friendly stimuli. Response time was calculated as the mean response time to correct go stimuli. Errors of omission, errors of commission and response time were calculated for the target rich single task, target rich dual task, target sparse single task, and target sparse dual task. The number of words recalled included the words heard through the headphones and the words participants said. Number of associated words was calculated as the number of times participants said an associated word aloud. An association was counted if the participant said a word back, regardless of what the word was (Darling & Helton, 2014). Number of associations and words recalled were calculated for the verbal recall task, target rich dual task, and target sparse dual task. The modified NASA-Task Load Index (NASA-TLX; Hart & Staveland, 1988) was administered by paper and pencil at the end of each of the five tasks, and was used to measure subjective workload on four dimensions (mental demand, physical demand, temporal demand, effort).

**Procedure**

Upon arrival, participants read an information form and completed a consent form. They were randomly assigned to a technology group, and provided with a demonstration on how to use the technology and given the opportunity to ask any questions (refer to Appendix C for the script used to give instructions). Those participants assigned to the Radio or Mobility App group held the iPhone, or had the radio clipped to them, for the duration of the experiment. The main body of the Radio was clipped to the participant’s pants. The microphone part of the Radio was clipped to the left lapel of their shirt if they were right-handed, or the right lapel if left-handed. Participants held the Taser in their dominant hand, and used their other hand to work the technology. Participants were instructed that all tasks would begin after hearing a high pitched tone followed by three low pitch tones, and the end of the task was signalled by a high pitched tone. All participants completed a practice verbal recall trial with four words to familiarise them with the technology, and to demonstrate the start and end tone. Feedback was provided about their performance, to ensure that they understood how to use the technology and followed task instructions.

Participants then completed practice trials for the target rich single task and target sparse single task, to familiarize them with the stimuli. They did not receive feedback on their performance in terms of commission or omission errors, but all participants were told at the end to ensure that they pointed at all stimuli, whether they were going to respond or not. Participants then completed the five experimental tasks. Lights were dimmed low for the duration of experiment. Balanced Latin squares was used to counterbalance order of conditions, and the order of word lists was also counterbalanced across conditions. Each task was 7 minutes and 37 seconds long. The background on the screens remained grey throughout experiment. Participants had a one minute seated break in between each task to ensure their arm did not get fatigued from holding the ‘Taser’. Participants completed the experiment one at a time. The total duration of the experiment was 70 minutes.

**Target rich and target sparse single tasks.** Participants were instructed to stand on the cross on the floor. Onscreen instructions then advised participants to point the Taser at all stimuli that appeared on the screen. This was done to ensure they visually searched for stimuli (as the large screens meant participants may not see stimuli), it mimics reality as Public Safety Officers have to be alert to all stimuli, and is in line with the method for the realistic firearm experiment (Wilson et al., 2015). Participants were instructed to respond to foes, who were holding guns, and to withhold responses to friendly people, who were not holding guns. They were also instructed to respond as fast and as accurately as possible. Following the original SART (Robertson et al., 1997), there were 225 trials. For each trial, the stimulus occurred in one of six locations (the left side or right side of one of the three screens). To ensure stimuli were shown in the six possible locations evenly, location of stimuli was quasi random (for 18 trials, each of the six possible locations would have three stimuli randomly appear). Each stimulus was presented for 1500 ms to enable the detection of the presence or absence of a gun (same as Wilson et al., 2015). This was followed by 500 ms interval with no stimulus. The total time from stimulus onset to next stimulus onset was 2000 ms. A response was counted as long as participants pushed the ‘Taser’ button before the next stimulus appeared. The stimuli were randomly sampled without replacement. For the target rich task, the percentage of foes was 89%, and in the target sparse task 11%. Participants heard scrambled words through headphones to control for auditory input received during the dual-task scenario. The two scrambled word lists were counterbalanced across the target rich and target sparse single tasks. Participants were instructed that they did not need to pay attention to these words. After each task, participants were instructed to return to the desk and complete the workload questionnaire.

**Verbal recall task.** Participants completed this task while standing in front of the screens on the white cross and looking at a grey background, and were required to still hold the “Taser” in their dominant hand. Participants were instructed to pay attention to the 20 words they would hear through the headphones and to radio (or say aloud if in control group) an associated word back every time they heard a word. They were also instructed they would need to recall as many words as possible (words heard and words said), at the end of the task. Requiring participants to say a word back enforces encoding of words (Green & Helton, 2011) while also reflecting the communicative nature of using the technology in reality. The use of the radio technology in this experiment was designed to reflect how the technology is used by Public Safety Officers. Therefore participants had to unlock the iPhone once at the beginning of the experiment to access the Mobility App before being able to radio a word back, and the Mobility App radio link would time out after 30 seconds if it was not used. Whist this does mean that the Radio is more easily accessible than the Mobility App, this reflects real-world constraints when using the technology. At the end of the task, participants were instructed that they have 130 seconds to return to the desk and recall as many words as possible. They were given a sheet with two columns marked on it- one for ‘words I heard’ and ‘words I said’. Words were marked as correct if they were recalled, regardless of what column they were written in. A Samsung Galaxy S6 was used to record words participants said, and was later transcribed to mark correct words recalled.

**Target rich and target sparse dual tasks.** In the two dual tasks, participants completed the target rich single task simultaneously with the verbal recall task, and the target sparse single task simultaneously with the verbal recall task. Participants were given the same instructions as the target rich and target sparse single tasks, the only difference being that they were also instructed they would hear 20 words rather than scrambled words. Participants were instructed that they needed to pay attention to the 20 words they would hear through the headphones and to radio an associated word back every time they heard a word. They were also instructed they would need to recall as many words as possible, both the words they heard and words they said, at the end of the task. The same stimuli was used for the target rich and target sparse dual tasks as the single tasks. Stimuli were randomly sampled without replacement, in a quasi-random location as with the single tasks. In the target rich dual task, 89% of stimuli were foes, and 11% were foes in the target sparse dual task. The audio tracks for the verbal recall tasks began at the same time as the target rich and target sparse dual task. At the end of the task, participants were instructed that they have 130 seconds to return to a desk and recall as many words as possible. No priority was given to either task in instructions to participants.

## Results

The dependent variables were calculated for each participant as: mean response time for correct responses for target rich and target sparse tasks, percentage of commission errors for target rich and target sparse tasks, percentage of omission errors for target rich and target sparse tasks, number of associations made in verbal recall tasks, number of words recalled in verbal recall tasks, and subjective workload ratings for each task.

Data was assessed for quality before running analyses. Outliers were assessed by looking at histograms and boxplots, and assessing whether there were scores that did not appear to belong to the distribution. Outliers were also assessed by determining whether any Z scores were greater than 3.29, as it is highly unlikely a score would fall more than three standard deviations from the mean, particularly in a small sample in the present study (Tabachnick & Fidell, 2013). Each dependent variable was assessed for each technology group separately. There were several outliers for commission and omission errors, in which several participants had a relatively high percentage of errors on one or more tasks compared to other participants. However, as this is a novel sustained attention task, these participants may have found the task more difficult than other participants, and one cannot assume that they did not listen to task instructions, particularly as no participant consistently scored higher across the four sustained attention tasks. One outlier was removed for physical demand rating in the verbal recall task, based on having a Z score higher than 3.29 and it also appeared distinct from the distribution on the histogram and box plot. Statistical tests were rerun with this outlier and results did not change. The one engineering student in the data set did not show up as an outlier in any dependent variables, so therefore was included in the analyses.

Dependent variables in each group were also assessed for normality by using the Shapiro-Wilk statistic and viewing histograms (Field, 2013). The majority of data was normally distributed. Commission and omission errors were positively skewed, indicating a possible floor effect. However, ANOVA’s are relatively robust to normality violations when group sizes are equal, as is the case in this study (Field, 2013). Please refer to Appendix D for descriptive statistics for all dependent variables.

**Response time**

Average response time for correct trials (correctly responded to a foe) was calculated for each participant for each task (target sparse and target rich tasks, both single and dual load). Response times for one participant did not record.

To test research questions 1a, 1b and hypothesis 1a, between-subjects correlations were conducted for each task, correlating errors (commission and omission) with the average response time to correct trials for the task. Between-subjects correlations removes within-subject variance and can reveal whether those that tend to respond faster make more errors for example, thereby revealing differences between people (Zelenski and Larsen 2000). These correlations were done collapsing groups, as these trends were not expected to differ between groups. Results are displayed in table 1.

Table 1.

*Correlations between response time and errors of omission and commission for the different tasks*

| | Commission errors | Omission errors |
|---|---|---|
| Target rich single | **-.31*** | **.28*** |
| Target sparse single | **-.30*** | .08 |
| Target rich dual | -.07 | **.46**** |
| Target sparse dual | .02 | .16 |

*Note.* * $p < .05$, ** $p < .001$

$N = 59$, apart from target sparse single where $N = 58$

There was a significant negative correlation between response time in the target rich single task and commission errors. This indicates there was a speed-accuracy trade-off in that those who responded faster were more likely to make errors of commission. As seen in Table 1, there was no speed-accuracy trade-off in the target rich dual task, as the correlation between response time and commission errors was non-significant. These results answer research question 1a. There was a significant positive correlation between omission errors and response time in the target rich single task and target rich dual task, which suggests those that took longer to respond were more likely to make errors of omission than those who were faster at responding. This result answers research question 1b. Response time and commission errors in the target sparse single task were significantly negatively correlated, which indicates a speed-accuracy trade-off. There were no significant correlations between response time and commission or omission errors in the target sparse dual task. This result partly supports hypothesis 1a.

To test hypothesis 1b, a 3 (technology group) x 2 (target probability; rich vs sparse) x 2 (load; single vs. dual) mixed ANOVA was conducted, with technology as a between-subjects variable. Response time was faster in the target rich tasks ($M$ = 1.034, $SE$ = .015) than the target sparse tasks ($M$ = 1.057, $SE$ = .012), $F(1, 55) = 7.26$, $p = .009$, $\eta_p^2 = .12$. This supports hypothesis 1b. Response time was also faster in the single tasks ($M$ = 1.030, $SE$ = .015) than the dual tasks ($M$ = 1.062, $SE$ = .013), $F(1, 55) = 13.64$, $p = .001$, $\eta_p^2 = .20$. There were no significant differences in response time between groups, $p > .05$.

As a side analysis, response time was broken down into six groups, based on the distance of the current stimuli from the previous stimuli. These groups were: 0 distance (current stimulus is in the same place as the previous stimulus), 1 (current stimulus is either directly to the left or right of previous stimulus), 2 from previous, 3 from previous, 4 from previous and 5 from previous. For all sustained attention tasks, response time increased as distance from previous target increased. Please refer to Appendix E for analysis of these linear trends.

**Commission errors**

For both the target rich and target sparse tasks, a commission error was when the participant incorrectly responded to friendly stimuli (in the target rich, there were 25 friendly stimuli, and 200 in the target sparse). This was calculated as the percentage of commission errors.

To tests hypotheses 2a, 2b, 2d and 2f, a 3 (technology group) x 2 (target probability; rich vs sparse) x 2 (load; single vs. dual) mixed ANOVA was conducted, with technology as a between-subjects variable. More errors of commission were committed in the target rich tasks (*M* = 10.17, *SE* = .98) than the target sparse tasks (M = .54, *SE* = .08), *F*(1, 57) = 100.71, $p < .001$, $\eta_p^2$ =.64. This supports hypothesis 2a. Errors of commission did not differ between the single tasks (*M* = 4.70, *SE* = .69) versus dual tasks (*M* = 6.01, *SE* = .51), *F*(1, 57) = 3.78, $p = .057$, $\eta_p^2$ =.06. This does not support hypothesis 2b. There were no significant group differences in the above trends, $p > .05$.

However at least one technology group had a different percentage of errors of commission overall than the other groups, *F*(2, 57) = 3.56, $p$ =.035, $\eta_p^2$ = .11. Preplanned orthogonal contrasts revealed no difference between the control condition (*M* = 4.61; *SE* =.87) versus using technology, $M_{difference}$ = -1.12 95% *CI* [-3.25; 1.01]. This does not support hypothesis 2f. However, as seen in Figure 5, those using the Mobility App (*M* = 4.23; *SE* = .87) made fewer commission errors than those using the Radio (*M* = 7.23; *SE* = .87), $p$ = .017, $M_{difference}$ = -3.01 95% *CI* [- 5.46; -.55]. This does not support hypothesis 2d.

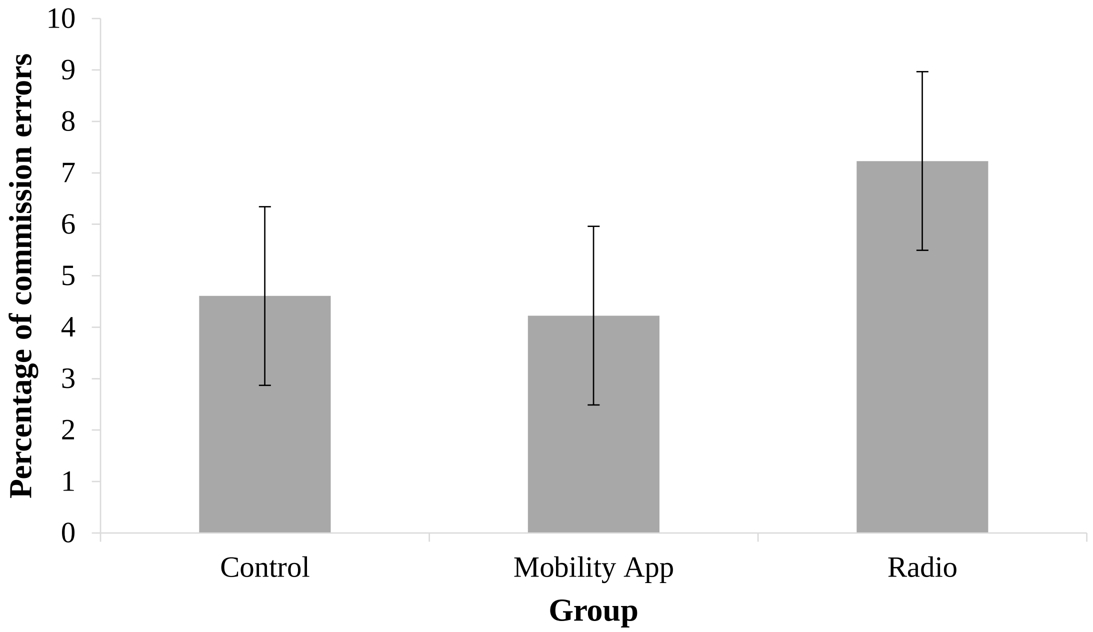


*Figure 5.* Overall mean percentage of commission errors for the different technology groups. Error bars are standard errors of the mean, 95% CI.

**Omission errors**

For both the target rich and target sparse tasks, an omission error is when the participant incorrectly withheld a response to foe stimuli (in the target rich there were 200 foe stimuli, and 25 in the target sparse). This was calculated as the percentage of omission errors.

To answer research question 2, and to test hypotheses 2c, 2e and 2f, a 3 (technology group) x 2 (target probability; rich vs sparse) x 2 (load; single vs. dual) mixed ANOVA was conducted, with technology group as a between-subjects variable. More omission errors were committed in the target sparse tasks ($M$ = 3.4, $SE$ = .43) than the target rich tasks ($M$ = 1.61, $SE$ = .26), $F(1, 57) = 19.13$, $p < .001$, $\eta_p^2 = .25$, which answers research question 2. More errors of omission were committed in the dual tasks ($M$ = 3.68, $SE$ = .50) than the single tasks ($M$ = 1.33, $SE$ = .25), $F(1, 57) = 20.06$, $p < .001$, $\eta_p^2 = .26$. This supports hypothesis 2c.

There were no group differences in omission errors, and no interaction effects involving groups, $p > .05$. This does not support hypothesis 2e or 2f.

**Word recall**

One participant was omitted from these analyses for failing to follow task instructions for recalling words. For the verbal recall task, target rich dual task and target sparse dual task, each participant was marked for how many words they correctly recalled (both words they heard and words they said aloud), out of a possible 40 words recalled. The number of associations made (out of 20) was also calculated for each participant. An association was counted if the participant said a word aloud. Descriptive statistics for associations made are displayed in Table 2.

Table 2.

Mean number of associations made in each task across the technology groups. Standard error is in brackets.

| | Recall only | Target Rich | Target Sparse |
|---|---|---|---|
| Control | 19.50 (.12) | 19.85 (.15) | 19.60 (.20) |
| Mobility App | 19.75 (.12) | 19.75 (.15) | 19.95 (.20) |
| Radio | 19.60 (.12) | 19.50 (.15) | 19.45(.20) |

A 3 (technology group) X 3 (task; recall only, target rich dual, target sparse dual) mixed ANOVA was conducted on associations made, with group as between subjects variable. Mauchly's test indicated the assumption of sphericity was not violated, $X^2(2) = .99$, $p = .93$, therefore sphericity was assumed. There was no difference in associations made between groups $F(2, 57) = 1.79$, $p = .18$, $\eta_p^2 = .06$, and no differences in associations made across tasks, $F(2, 114) = .53$, $p = .59$, $\eta_p^2 = .01$.

To test hypothesis 3a, 3b, and 3c a 3 (technology group) X 3 (task; recall only, target rich dual, target sparse dual) mixed ANOVA was conducted on number of words recalled, with groups as a between subjects variable. Mauchly's test indicated the assumption of sphericity was not violated, $X^2(2) = .96$, $p = .30$, therefore sphericity was assumed. There was a main effect for task, $F(2, 112) = 15.97$, $p < .001$, $\eta_p^2 = .22$. Preplanned orthogonal contrasts revealed more words were recalled in the recall only task ($M = 17.32$, $SE = .66$), than the dual tasks, $F(1, 56) = 27.60$, $p < .001$, $\eta_p^2 = .33$. This supports hypothesis 3a. There was no difference in words recalled between the target sparse dual task ($M = 14.64$, $SE = .53$) and target rich dual task ($M = 14.11$, $SE = .50$), $F(1, 56) = .86$, $p = .36$, $\eta_p^2 = .02$.

At least one technology group had a different number of words recalled overall than the other groups, $F(2, 56) = 4.73$, $p = .013$, $\eta_p^2 = .14$. Preplanned orthogonal contrasts were conducted, first comparing technology (Mobility App and Radio) to the control group ($M = 17.28$; $SE = .77$), and then comparing the Radio ($M = 14.51$, $SE = .79$) and Mobility App ($M = 14.28$, $SE = .77$). More words were recalled in the control condition compared to using technology, $p = .003$, $M_{difference} = 2.89$ 95% $CI$ [1.00; 4.78]. This result supports hypothesis 3c. However, there was no difference in words recalled between the Mobility App and Radio, $p = .84$, $M_{difference} = -.23$ 95% $CI$ [- 2.42; 1.97], which does not support hypothesis 3b.

There was a significant interaction between task and group, $F(4, 112) = 5.42$, $p < .001$, $\eta_p^2 = .16$. As seen in Figure 6, the trend of fewer words recalled in the dual tasks compared to verbal recall task is only the case for the Radio and control. In fact, the Mobility App appears to have somewhat low word recall across all three tasks, which somewhat supports hypothesis 3b.

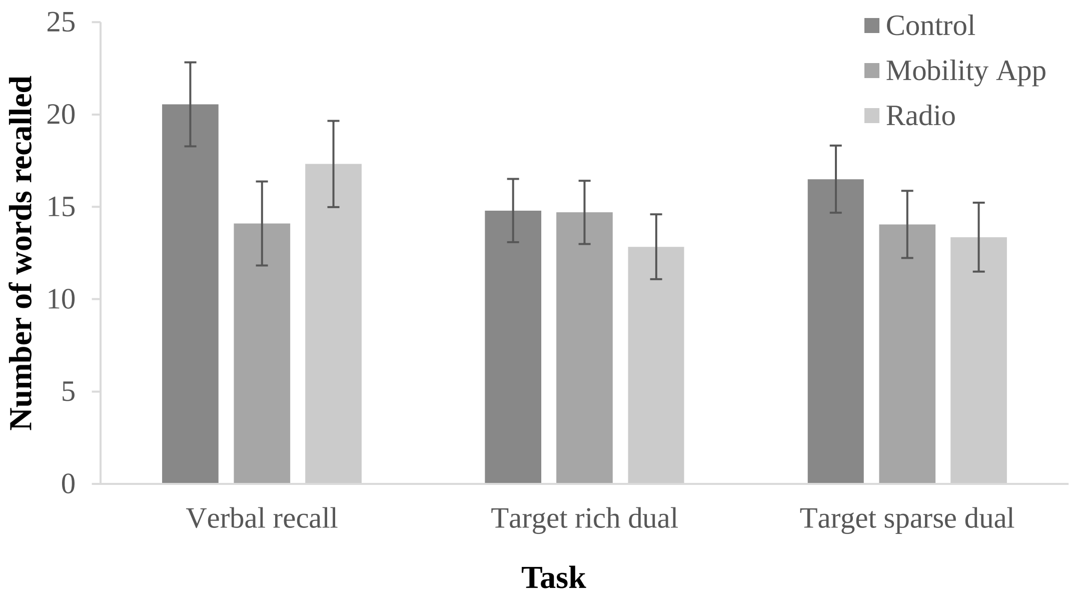


*Figure 6.* Mean number of words recalled across the three different tasks for each technology group. Error bars are standard errors of the mean, 95% CI

**Workload**

For each participant, workload ratings for each task for each subjective workload measure were scored. To test hypotheses 4a, 4b, 4c, and 4d, a 3 (technology) x 2 (target probability; rich vs sparse) x 2 (load; single vs. dual) mixed ANOVA was conducted for each workload measure, with technology group as a between-subjects variable.

Dual tasks (*M* = 72.11 *SE* = 2.09) were rated as more mentally demanding that single tasks (*M* = 45.95, *SE* = 2.71), $F(1, 57) = 122.33$, $p < .001$, $\eta_p^2 = .68$. Dual tasks (*M* = 43.73, *SE* = 3.05) were rated as more physically demanding than single tasks (*M* = 36.43, *SE* = 2.86), $F(1, 57) = 12.25$, $p = .001$, $\eta_p^2 = .18$. Dual tasks (*M* = 62.25, *SE* = 2.25) were rated as more temporally demanding than the single tasks (*M* = 46.62, *SE* = 2.59), $F(1, 57) = 51.58$, $p < .001$, $\eta_p^2 = .48$. Dual tasks (*M* = 71.75, *SE* = 1.99) were rated as more effortful than the single tasks (*M* = 50.88, *SE* = 2.56), $F(1, 57) = 90.89$, $p < .001$, $\eta_p^2 = .62$. These results support hypothesis 4b.

There were no significant differences in mental demand ratings for target rich task ($M$ = 60.03, $SE$ = 2.24) and target sparse tasks ($M$ = 58.04, $SE$ = 2.28), $p > .05$. Target rich tasks ($M$ = 41.51, $SE$ = 2.86) were rated as more physically demanding than target sparse tasks ($M$ = 38.65, $SE$ = 2.85), $F(1, 57) = 4.21$, $p = .045$, $\eta_p^2 = .07$. Target rich tasks were rated as more temporally demanding ($M$ = 56.54, $SE$ = 2.21) than target sparse tasks ($M$ = 52.33, $SE$ = 2.39), $F(1, 57) = 7.31$, $p =.009$, $\eta_p^2 = .11$. Target rich tasks were rated as more effortful ($M$ = 62.97, $SE$ = 2.11) than target sparse tasks ($M$ = 59.66, $SE$ = 2.39), $F(1, 57) = 5.32$ , $p =.025$, $\eta_p^2 = .09$. These results support hypothesis 4a, excluding mental demand ratings which does not support hypothesis 4a.

There were no other significant interaction effects, or overall differences between groups, $p > .05$. This does not support hypotheses 4c and 4d.

## Discussion

This study aimed to explore high-go/low no-go performance in a more ecologically valid paradigm, and to compare two forms of communication technology used by Public Safety agencies. The results will each be discussed in turn, and related to hypotheses and past research.

### Response time

The current target rich tasks appeared to bias participants towards fast responding compared to the target sparse tasks. Response time was faster in the target rich tasks than the target sparse task, which supports hypothesis 1b. Consistent with previous research, the current high-go/low no-go task biased participants towards fast responding compared to the perceptually equivalent low-go vigilance task (Dippel, Chmielewski, Muckschel, & Beste, 2015; Helton, & Russell, 2011; Helton, Weil, Middlemiss, & Sawers, 2010; Stevenson, Russell, & Helton, 2011). This is in line with the response inhibition perspective of high-go

SART tasks, which argues that the SART biases participants towards fast responding. This result does not support mindlessness perspective, which considers both the high-go SART and low-go vigilance tasks to be measures of sustained attention. The mindlessness perspective would argue that the faster response times in the target rich tasks compared to target sparse tasks was due to the perceptual features of altering task format, which is not entirely plausible as the same stimuli were used for all tasks (Stevenson, Russell, & Helton 2011).

Response times were faster in the single tasks than the dual tasks. Reaction times have been shown to be slower in a dual high-go task than a single high-go task (Grandjean & Collette, 2011), whereas as other studies have found no difference in response time when a secondary task is added to a high go/low no-go task (Head & Helton, 2014), which differs from the result in the current study. The current study required participants to physically move and subdue foe threats, and this differs to past studies which have used computer based measures of sustained attention. Therefore the difference in findings may be attributed to the increased physicality required in the current sustained attention task, and therefore when a secondary task is added, response time slows.

In answer to research question 1a, a speed-accuracy trade-off was found in the target rich single task but not in the target rich dual task. For the target rich single task, commission errors were negatively correlated with response time (speed-accuracy trade-off), indicating that those who responded faster were more likely to make commission errors. This was not the case for the target rich dual task. A speed-accuracy trade-off is a key part of the traditional computer SART, as participants develop a prepotent motor response that is difficult to inhibit and accuracy is sacrificed for fast responding (Helton, Weil, Middlemiss, & Sawers, 2010; Stevenson, Russell, & Helton, 2011; Wilson et al., 2015). Forcing participants to slow their responses should reduce the speed-accuracy trade-off (Head &

Helton, 2014; Seli, Jonker, Solman, Cheyne, & Smilek, 2013). In the more realistic firearm SART, a speed-accuracy trade-off was found, although the present study differs in that stimuli location was unpredictable (Wilson et al., 2015). Manual selection and unpredictable location of stimuli has been shown to eliminate the speed-accuracy trade-off by forcing participants to slow down (Helton & Head, 2014), and the present study required manual selection of stimuli and location was unpredictable. The correlation between response time and commission errors in the present study for the target rich single task ($r$ = -.31) was weaker than previous studies, which have ranged from $r$ = -.54 (Head & Helton, 2013) to $r$ = -.64 (Wilson et al., 2015). This suggests that unpredictable location and manual selection in the present study may have reduced the speed-accuracy trade-off to some extent, although it did not eliminate it in the target rich single task. The lack of speed-accuracy trade-off in the dual task suggests that there is something about the dual task that eliminates the speed-accuracy trade-off. The increased physicality and additional cognitive load of also completing a verbal recall task in the dual task may have slowed participants down sufficiently to reduce the speed-accuracy trade-off. This is supported by the result that response times were slower in dual tasks than single tasks. There may be a critical time in which responses must be delayed to eliminate or reduce the speed-accuracy trade-off, as compared to 800ms suggested in past studies (Seli, Jonker, Solman, Cheyne, & Smilek, 2013).

There was a speed-accuracy trade-off in the target sparse single condition, but not in the target sparse dual condition. This result only partially supports hypothesis 1a, as it was predicted there would be no speed-accuracy trade-offs in the target sparse tasks, as previous studies have shown that low-go/high no-go tasks do not bias participants towards fast responding (Helton, Weil, Middlemiss, & Sawers, 2010; Stevenson, Russell, & Helton 2011). Perhaps a speed-accuracy trade-off developed in both the single tasks (target sparse *and* target rich) due to the nature of the stimuli. For example, weapon stimuli are known to cause aggressive responding in participants (Berkowitz & LePage, 1967), and this may have manifested in the current study as participants responding fast to all stimuli. However, there were no speed-accuracy trade-offs in the dual tasks even though the same stimuli were used. Commission errors did not differ between dual tasks and single tasks, but response time slowed in dual tasks compared to single tasks. This slowing in response time may have been enough to reduce the speed-accuracy trade-off in the dual tasks. The possible reasons why responses were slower in dual tasks compared to single tasks has been discussed above. Future studies should use number stimuli utilizing the Vision Space and 'Taser' to investigate if the trend of speed-accuracy trade-offs in the single tasks is influenced by stimuli holding guns, or due to a more realistic task in general that involves physical movement. Another option would be to use the same people stimuli as the present study, but instead foes are indicated by a non-weapon (for example, holding a cup). These studies could determine whether it is the weapon stimuli in particular that influences speed-accuracy trade-off in single sustained attention tasks, or whether it is influenced by a more realistic and physical task. This is critical to determine, as Police Officers for example experience threats every day while having to sustain attention, and this may influence their ability to sustain their attention.

Research question 1b aimed at determining whether there was a relationship between response time and errors of omission. Response time was positively correlated with omission errors in the target rich single task and target rich dual task. Participants who responded slower in both these tasks were more likely to make errors of omission by failing to respond to foe stimuli, and this correlation was stronger for the dual task than single. There were no significant correlations between omission errors and response time in either of the target sparse conditions. Previous studies have shown that unpredictable location and manual stimuli selection leads to a positive correlation between errors of omission and response time, and the authors suggested that participants found manual selection difficult and missed responding to ‘go’ stimuli due to the additional cognitive load of having to search for stimuli (Head & Helton, 2013, 2014). In the present study, those participants who took longer to respond overall in the target rich tasks may have run out of time to respond or may have failed to see the stimuli. Participants also rated the target rich tasks as having more subjective workload than the target sparse conditions. Therefore participants may have slowed down as they perceived the target rich tasks as more subjectively difficult.

**Errors of commission and omission**

More errors of commission were committed in the target rich tasks than the target sparse tasks, which supports hypothesis 2a and the response inhibition perspective of high-go tasks. Previous studies have found this result, and argued that target rich tasks introduce response inhibition demands, and therefore participants are more likely to incorrectly respond to stimuli, whereas the traditional target sparse vigilance tasks do not introduce inhibition demands (Carter, Russell, & Helton, 2013; Dippel, Chmielewski, Muckschel, & Beste, 2015; Helton, Weil, Middlemiss, & Sawers, 2010; Helton & Russell, 2011; Wilson et al., 2015). The current sustained attention tasks had perceptually equivalent stimuli, the only difference was that target rich tasks were hypothesized to require inhibiting a prepotent motor response whereas the target sparse tasks do not allow the opportunity for a prepotent motor response to develop (Carter, Russell, & Helton, 2013). Therefore the increased errors of commission in the present study’s target rich tasks compared to target sparse tasks can be attributed to prepotent motor response in target rich task, thereby supporting response inhibition perspective of high-go/low no-go tasks. This result also does not support mindlessness theory, which does not take into account response inhibition and considers that perceptually equivalent target rich high-go tasks and target sparse low no-go tasks are both measures of sustained attention. Mindlessness would argue that any performance differences are attributable to the perceptual features of altering task format, which is not entirely plausible as the same stimuli were used for all tasks (Carter, Russell, & Helton, 2013; Stevenson, Russell, & Helton 2011).

Errors of commission did not differ between the single versus dual tasks. This does not support hypothesis 2b. There were no interaction effects, meaning this trend was the same for target rich and target sparse tasks. This differs to previous research which has found difficult secondary tasks, similar to the secondary task used in the present study in which multiple processing stages are used, disrupt one’s ability to inhibit a prepotent response. This was evidenced by an increase in commission errors on dual task compared to single (Grandjean & Collette, 2011; Head & Helton, 2014). In the present study, participants reported higher subjective workload ratings for the dual tasks than single tasks, suggesting that the dual tasks did in fact place additional load on participants, although this did not translate into worse performance in terms of commission errors. There is the possibility that the recall task did not place sufficient cognitive load on participants to disrupt their ability to inhibit responses in the target rich task, and sustain their attention in target sparse task (Head & Helton, 2014). However, mindlessness theory is also not supported as there was no increase in performance from the single to dual tasks. This suggests that providing more stimuli in the dual tasks does not necessarily capture exogenous attention and reorient attention back to the sustained attention task (Head & Helton, 2014; Manly et al., 2004).

More errors of omission were committed in the target sparse tasks than the target rich tasks which answers research question 2. Most studies have found no difference in omission errors between perceptually equivalent high-go tasks and low-go tasks (Carter, Russell, Helton, 2013; Dippel, Chmielewski, Muckschel, & Beste, 2015; Helton & Russell, 2011), although high-go tasks have been shown to have more omission errors than perceptually equivalent low-go tasks (Helton, Weil, Middlemiss, & Sawers, 2010). The response inhibition perspective of high-go SART performance focuses on commission errors, and it has been suggested omission errors are more likely to reflect actual lapses in sustained attention (Helton, Weil, Middlemiss, & Sawers, 2010). However, high-go/low no-go tasks confound sustained attention with response inhibition, making it difficult to get a pure measure of sustained attention.

More errors of omission were committed in the dual tasks than the single tasks. This supports hypothesis 2c, and supports previous research (Doneva & de Fockert, 2014; Head & Helton, 2014; Wilson et al., 2015). Particularly in reference to high-go/low no-go tasks, omission errors are interpreted as ‘breathers’ or strategic pauses, or participants taking a break from the task due to the additional cognitive load of having to complete two tasks at once compared to a single task (Helton, Head, & Russell, 2011). As mentioned previously, dual tasks were rated as subjectively having more workload than the single tasks, and dual tasks also had a slower response time than single tasks, indicating that participants may slow down when a task is hard. This evidence lends itself to the idea that omission errors may be breathers.

There were minimal group differences in technology used. There were no differences in commission or omission errors for using technology versus not using any technology, which does not support hypothesis 2f. Participants in the control group were hypothesized to make the fewest errors of omission and commission as they did not need to interact with technology that could have disrupted their performance, although the results do not support this hypothesis. Participants in the Radio condition overall made more errors of commission than those using the Mobility App, which does not support hypothesis 2d. This result does not conform to theory, as the Mobility App was hypothesized to place more cognitive load on participants (as it involved checking the interface and ensuring the radio connection remained established), whereas the Radio does not require participants to look away from the main screens or monitor connection. There were no overall group differences in omission errors, which does not support hypothesis 2e. The verbal recall task or sustained attention tasks may have required increased difficulty to elicit group differences. Previous research has shown that participants make commission errors 25% to 50% of the time (Doneva & De Fockert, 2014), whereas in the current study, participants made commission errors 9.07% in the target rich single task and 11.27% in the target rich dual task which is considerably lower than previous studies (e.g. Dippel, Chmielewski, Muckschel, & Beste, 2015; Head & Helton, 2013; Wilson et al., 2015). There was also a positive skew in data, with a number of participants committing no errors of commission or omission. This evidence suggests that the current modified high go/low no-go task may be relatively easy compared to previous traditional sustained attention tasks. Therefore any differences between technologies may be elicited if task difficulty is increased.

**Word recall and workload ratings**

More words were recalled in the verbal recall task than the dual sustained attention tasks (target rich and target sparse), which supports hypothesis 3a. This supports past research which has used the same word lists, even though the pace of words was slower in the present study (Head & Helton, 2014). This also supports past research which has investigated word recall and dual task performance in physically demanding tasks such as rock climbing (Darling & Helton, 2014; Green & Helton, 2011). There was no difference in word recall between the target rich dual task and target sparse dual task. The finding that more words were recalled in the verbal recall task than the dual sustained attention tasks appeared to only be true for those in the Radio and control groups. Participants using the Mobility App had somewhat low word recall across all three tasks. This suggests that word recall performance with the Mobility App is not impacted with the addition of a secondary task, as performance is already relatively low. This somewhat supports hypothesis 3b.

More words were recalled by participants using no technology compared to participants who were using technology, which supports hypothesis 3c. However, there were no group differences in subjective workload ratings, which does not support hypothesis 4c and 4d. This suggests that although participants did not differ in their perception of workload based on what technology they were using, technology somewhat affected their performance. Those participants that used technology had to not only recall words, but also had to interact with technology, whereas those participants using no technology did not have the additional cognitive load of having to interact with technology. There was no difference in words recalled overall between the Mobility App and Radio, which does not support hypothesis 3b. The Mobility App was hypothesized to place more cognitive load on participants as it involved checking the interface and ensuring the radio connection remained established. Although there was no difference in word recall overall for type of technology used (Radio versus Mobility App), the impact of adding a secondary task on word recall differs for these two technologies, as mentioned previously.

Target rich tasks were rated as having more subjective workload than target sparse tasks. This result supports hypothesis 4a. This result was found in the realistic firearm SART, and the authors suggested that the high-go SART requires more cognitive resources than a low-go vigilance task as SART also requires participants to inhibit a prepotent motor response, whereas low-go vigilance tasks do not (Wilson et al., 2015). The result in the present study therefore supports the response inhibition perspective. The result does not support mindlessness theory which would predict no difference in workload ratings as traditional high-go SART and low-go vigilance task with perceptually equivalent stimuli are both considered measures of sustained attention. The dual task was rated as having more subjective workload than the single task. This result supports hypothesis 4b. The dual tasks were hypothesized to add additional cognitive load to the single tasks, and therefore had higher subjective workload ratings.

**Implications**

These results contribute to existing theory by supporting past research that has argued response inhibition plays a part in high-go/low no-go sustained attention tasks. The response inhibition perspective still appears relevant to high-go tasks were stimuli are novel and dynamic (each trial had different people stimulus). Unpredictable stimuli location and manual selection did not remove the speed-accuracy trade-off when cognitive load was lower (single task), although it may have reduced the speed-accuracy trade-off in comparison to previous studies. There was an unexpected result of a speed-accuracy trade-off in the target sparse condition. While this could be experimental noise, it could possibly be due to the introduction of novel stimuli. This implies that researchers need to consider how the introduction of novel stimuli could bias participant’s responses, independent of sustained attention ability or response inhibition. Introducing novel stimuli could also influence participants’ ability to inhibit a prepotent response, particularly as previous studies show that anxiety for example can aid response inhibition (Wilson, Russell, & Helton, 2015).

The current results have implications for the real-world. The dual tasks in the current study can be likened to a Police Officer radioing dispatch while searching for an offender. Based on the current results, the Officer would fail to find the offender more often than if he were to be doing each task one at a time, but incorrectly responding to a neutral passer-by would be unaffected by the addition of a secondary task. Contrary to mindlessness, adding a secondary task does not necessarily improve performance, and can in most cases be detrimental to performance. This suggests that adding exogenous cues to capture attention will actually interfere with sustained attention performance and decrease ability to remember information. For example, if a Police officer is listening to information over the Radio, whilst completing a dangerous task that requires sustained attention, information should be presented in a way which minimally interferes with the attention task. There was some indication that the Mobility App interferes with ability to recall words, however further studies are required to elucidate the difference between communication technologies.

**Methodological considerations and future directions**

One possible limitation for the present study is the type of participants sampled. Participants were first year Psychology students, the majority of them female. This was done due to time and resource constraints, as the experiment was relatively long in duration and it was difficult to recruit participants from other areas and reward them within budget constraints. SART performance has been shown not to differ across gender, age and education (Chan, 2001), and undergraduate students are traditionally recruited as participants for sustained attention studies (e.g.Finkbeiner, Wilson, Russell, & Helton, 2015; Seli, Jonker, Cheyne, & Smilek, 2013; Wilson, Russell, & Helton, 2015). Although previous studies have shown females report higher scores of subjective workload on NASA-TLX than males (Hancock, 1989), the current study found no gender differences in ratings of subjective workload. Therefore this is not a severe limitation. While the current study has tested the novel high-go/low no-go task with these participants to explore whether this modified task extends past research, future studies could explore SART/high-go performance using Police Officers or those trained to deal with serious sustained attention tasks. For example, Police Officers are familiar with radio communication technology and often experience threatening situations in which high demands are placed on sustained attention.

Another possible limitation in the present study is that several useful performance metrics were not obtained. Due to the nature of the task, response times before and after both commission and omission errors were not calculated, as distance from previous target stimuli varied greatly across all response times and the error rate was relatively low overall. This problem could have been solved by calculating participant's reaction time for each correct go trial, in terms of when they first start moving the 'Taser' when stimuli first appear. This was attempted in the present study, but there was no clear reaction time as participants appeared to move the "Taser" continuously. In previous research these metrics have been used to add strength to the response inhibition theory and show that participants cycle between strategies (Helton, Head, & Russell, 2011; Manly et al., 2000). Likewise, tracking head and 'Taser' movements would have provided interesting data. However, the present study provides valuable results that contribute towards the debate in arguing that response inhibition plays a part in the high go/low no-go sustained attention tasks.

Another limitation is that the race of people stimuli could have biased participants responding. A vigilance paradigm using pictures of African American or White people, some holding a gun and others neutral objects, found that participants were more likely to incorrectly shoot African American people than White people (Greenwald, Oakes, & Hoffman, 2003). The racial bias of being more likely to shoot non-whites is common in the literature (e.g. Correll, Park, Judd, Wittenbrink, Sadler, & Keesee, 2007; Unkelbach, Forgas, & Denson, 2008). Although not within the scope of the current study, it is possible that participants were biased to responding towards a certain race (for example, shooting non-whites whether they were holding a gun or not). However, people stimuli were carefully designed to ensure an even number of each race, and a variation of faces and races were randomly assigned to foe and friendly. Therefore the possibility of this influencing results was mitigated, and unlikely to have affected results.

Future studies could increase the difficulty of the present sustained attention tasks, as it is possible that the present modified high-go/low no-go task was easier than the traditional SART. Past experiments often have varying size of stimuli for each trial to ensure participants are not responding to only perceptual features of stimuli (e.g. Helton, Weil, Middlemiss, & Sawers, 2010). It is possible that participants responded to perceptual features of the gun in the current experiment as gun size was not varied, when in reality guns come in a variety of sizes and shapes. Decreasing the contrast between stimuli and background has been shown to decrease performance in traditional vigilance tasks (Helton and Warm, 2008). Increasing the rate of stimuli presentation may make the current high-go/low no-go task more difficult, as previous studies present stimuli around 250ms (Head & Helton, 2014; Helton, Head, Russell, 2011; Robertson et al., 1997; Seli, Cheyne, & Smilek, 2012). However, the current study has the same target salience and pace as the more ecologically valid firearm emitter SART (Wilson et al., 2015), and a faster pace of presentation would likely increase the rate of omission errors irrespective of sustained attention, as participants would not be physically capable of turning around fast enough to perceive stimuli. It is also not necessarily realistic that stimuli in the real world would appear at a pace of 250ms, and therefore increasing pace of stimuli presentation may not be justifiable from an ecological perspective. Future studies should create more difficult sustained attention tasks that reflect real world constraints. For example, it is justifiable to decrease the salience of weapons (as people may have them hidden in reality) or make target stimuli vary (foes could hold a variety of weapons rather than the same gun). Alternatively, including background noise (the radio or static noise), could reflect issues of sustained attention when someone is driving and watching the road for dangers. Increasing task difficulty may also elicit differences between technology groups better. An alternative explanation to participants finding the current hi-go/low no-go task easier than traditional SART's is that people perform better in tasks that reflect real world constraints compared to computer-based SART.

While not within the scope of the current study, paradigms should be created that differentiate between active learners versus passive perceptual learners (Helton & Head, 2015). Helton and Head (2015) argue that introducing novel stimuli may activate perceptual processes that confound measures of sustained attention. Passive perceptual learners are those who over time learn the stimuli and become better at the task, and therefore performance increases with time on task, contrary to traditional sustained attention tasks. Active learners however are more proactive in learning the novel stimuli and perform well to begin with, but performance decreases over time which is typical in a vigilance experiment. Engagement in these processes depends on individual differences. The authors completed a study in which participants that correctly identified the first target stimuli were classified as active searchers, and other participants as passive perceptual learners. The classification of participants into passive perceptual learners and active searchers was a limitation (Head & Helton, 2015). Future studies should investigate whether other perceptual processes are involved in sustained attention tasks which use novel and dynamic stimuli, as this may confound interpretation of results. This would require including periods of watch (as compared to running each task continuously as was done in the present study), and determining how to classify active searchers and passive perceptual learners.

## Concluding remarks

Overall, the novel sustained attention tasks in the current study extend previous research that argues high-go/low no-go tasks measure response inhibition. Results indicated that the introduction of novel stimuli may have biased fast responding irrespective of sustained attention ability, particularly in terms of the speed-accuracy trade-off, and this is an important area for future research. The current results have implications for the real-world, particularly for Public Safety Officers who sustain attention in dangerous situations while also using radio communication technology. Future studies should assess sustained attention performance in more ecologically valid settings to not only test theories of sustained attention, but to determine what other cognitive processes are involved in high-demand stressful situations.

## References

Ababsa, F., Didier, J. Y., Tazi, A., & Mallem, M. (2007). Software architecture and calibration framework for hybrid optical IR and vision tracking system. In *Control & Automation, 2007. MED'07. Mediterranean Conference on* (pp. 1-6). IEEE.

Bellgrove, M. A., Hawi, Z., Kirley, A., Gill, M., & Robertson, I. H. (2005). Dissecting the attention deficit hyperactivity disorder (ADHD) phenotype: sustained attention, response variability and spatial attentional asymmetries in relation to dopamine transporter (DAT1) genotype. *Neuropsychologia*, *43*(13), 1847-1857.

Berkowitz, L., & LePage, A. (1967). Weapons as aggression-eliciting stimuli. *Journal of Personality and Social Psychology*, *7*(2), 202-207.

Carter, L., Russell, P. N., & Helton, W. S. (2013). Target predictability, sustained attention, and response inhibition. *Brain and Cognition, 82*(1), 35-42.

Chan, R. C. (2001). A further study on the sustained attention response to task (SART): The effect of age, gender and education. *Brain Injury*, *15*(9), 819-829.

Cheyne, J.A., Carriere, J. S. A., & Smilek, D. (2009). Absent minds and absent agents: Attention-lapse induced alienation of agency. *Consciousness and Cognition*, *15,* 578-592.

Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). Across the thin blue line: police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology*, *92*(6), 1006-1023.

Darling, K.A., & Helton, W.S. (2014). Dual-task interference between climbing and a simulated communication task. *Experimental Brain Research, 232*(4), 1367-1377.

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance.* Academic Press, London

Dippel, G., Chmielewski, W., Mückschel, M., & Beste, C. (2015). Response mode-dependent differences in neurofunctional networks during response inhibition: an EEG-beamforming study. *Brain Structure and Function, 221,* 1-11.

Dockree, P.M., Bellgrove, M.A., O'Keefe, F.M., Moloney, P., Aimola, L., Carton, S., & Robertson, I.H. (2006). Sustained attention in traumatic brain injury (tbi) and healthy controls: enhanced sensitivity with dual-task load. *Experimental Brain Research, 168*(1), 218-229.

Dockree, P.M., Kelly, S.P., Roche, R.A.P., Hogan, M.J., Reilly, R.B., Robertson, I.A. (2004). Behavioural and physiological impairments of sustained attention after traumatic brain injury. *Cognitive Brain Research, 20*(3), 403-414.

Doneva, S. P., & De Fockert, J. W. (2014). More conservative go/no-go response criterion under high working memory load. *Journal of Cognitive Psychology, 26*(1), 110-117.

Farrin, L., Hull, L., Unwin, C., Wykes, T., & David, A. (2003). Effects of depressed mood on objective and subjective measures of attention. *The Journal of Neuropsychiatry and Clinical Neurosciences, 15,* 98-104.

Field, A. (2013). *Discovering statistics using SPSS*. Sage publications. Fourth ed.

Finkbeiner, K. M., Wilson, K. M., Russell, P. N., & Helton, W. S. (2015). The effects of warning cues and attention-capturing stimuli on the sustained attention to response task. *Experimental Brain Research, 233*(4), 1061-1068.

Fronczek, R., Middelkoop, H. A., Van Dijk, J. G., & Lammers, G. J. (2006). Focusing on vigilance instead of sleepiness in the assessment of narcolepsy: high sensitivity of the Sustained Attention to Response Task (SART). *Sleep, 29*(2), 187-191.

Grandjean, J., & Collette, F. (2011). Influence of response prepotency strength, general working memory resources, and specific working memory load on the ability to inhibit predominant responses: a comparison of young and elderly participants. *Brain and Cognition, 77*(2), 237-247.

Green, A.L., & Helton, W.s. (2011). Dual-task performance during a climbing traverse. *Experimental Brain Research, 215,* 307-313.

Greenwald, A. G., Oakes, M. A., & Hoffman, H. G. (2003). Targets of discrimination: Effects of race on responses to weapons holders. *Journal of Experimental Social Psychology*, *39*(4), 399-405.

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., & Parasuraman, R. (2003). The vigilance decrement reflects limitations in effortful attention, not mindlessness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(3), 349-359.

Hancock, P. A. (1989). The effect of performance failure and task demand on the perception of mental workload. *Applied Ergonomics, 20*(3), 197-205.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology, 52*, 139-183.

Head, J., & Helton, W. S. (2012). Natural scene stimuli and lapses of sustained attention. *Consciousness and Cognition, 21*(4), 1617-1625

Head, J., & Helton, W.S. (2013). Perceptual decoupling or motor decoupling? *Consciousness and Cognition, 22,* 913-919.

Head, J., & Helton, W. S. (2014). Sustained attention failures are primarily due to sustained cognitive load not task monotony. *Acta Psychologica, 153,* 87-94.

Head, J., & Helton, W.S. (2015). Passive perceptual learning versus active searching in a novel stimuli vigilance task. *Experimental Brain Research, 233,* 1481-1489.

Helgason, D., N. Francis, and J. Ante, Unity3d. 2005, Unity Technologies.

Helton, W. S. (2009). Impulsive responding and the sustained attention to response task. *Journal of Clinical and Experimental Neuropsychology*, *31*(1), 39-47.

Helton, W.S., Head, J., & Russell, P.N. (2011). Reliable- and unrelaible-warning cues in Sustained Attention to Response Task. *Experimental Brain Research, 209*¸401-407.

Helton, W. S., Kern, R. P., & Walker, D. R. (2009). Conscious thought and the sustained attention to response task. *Consciousness and Cognition*, *18*(3), 600-607.

Helton, W.S., & Russell, P.N. (2011). Feature absence-presence and two theories of lapses of sustained attention. *Psychological Research, 75,* 384-392.

Helton, W.S., & Warm, J.S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica, 129,* 18-25.

Helton, W.S., Weil, L., Middlemiss, A., & Sawers, A. (2010). Global interference and spatial uncertainty in the Sustained Attention to Response Task (SART). *Consciousness and Cognition, 19*(1), 77-85.

Knowles, D., & Tay, R. (2002). Driver inattention: More risky than the fatal four? In Proceedings of the 2002 Road Safety Research, Policing and Education Conference (pp. 377–392). Adelaide, South Australia: Causal Productions.

Mackworth, N. H. (1948). The breakdown of vigilance durning prolonged visual search. *Quarterly Journal of Experimental Psychology, 1*(1), 6-21,

Manly, T., Davison, B., Heutink, J., Galloway, M., & Robertson, I. (2000). Not enough time or not enough attention?: Speed, error and self-maintained control in the Sustained Attention to Response Test (SART). *Clinical Neuropsychological Assessment*, *3*, 167-177.)

Manly, T., Heutink, J., Davison, B., Gaynord, B., Greenfield, E., Parr, A., ... & Robertson, I. H. (2004). An electronic knot in the handkerchief: “Content free cueing” and the maintenance of attentive control. *Neuropsychological Rehabilitation*, *14*(1-2), 89-116.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: further investigations of sustained attention to response. *Neuropsychologia*, *37,* 661-670.

Paivio, A., Yuille, J.C. & Madigan S.A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76*, 21–25.

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science, 205*(4409), 924-927.

Peebles D., & Bothell, D. (2004). Modelling performance in the sustained attention to response task. In *ICCM*, 231–236. Carnegie Mellon University/University of Pittsburgh, Pittsburgh

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*(6), 747-758.

Seli, P., Cheyne, J.A., & Smilek, D. (2012). Attention failures versus misplaced diligence: Separating attention lapses from speed-accuracy trade-offs. *Consciousness and Cognition*, *21*, 277-291.

Seli, P., Jonker, T.R., Solman, G.J.F., Cheyne, J.A., & Smilek, D. (2013). A methodological note on evalutating performance in a sustained-attention-to-response task. *Behavior Research Methods, 45*(2), 355-363.

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition*, *13*(4), 657-690.

Smallwood, J., & Schooler, J.Q. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946-958.

Smilek, D., Carriere, J. S., & Cheyne, J. A. (2010). Failures of sustained attention in life, lab, and brain: ecological validity of the SART. *Neuropsychologia*, *48*(9), 2564-2570.

Stevenson, H., Russell, P. N., & Helton, W. S. (2011). Search asymmetry, sustained attention, and response inhibition. *Brain and Cognition*, *77*(2), 215-222.

Tabachnick, B. G., & Fidell, L. S. (2013). Using Multivariate Statistics, 6th International edition (cover) edn.

Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *42*(2), 183-194.

Unkelbach, C., Forgas, J. P., & Denson, T. F. (2008). The turban effect: The influence of Muslim headgear and induced affect on aggressive responses in the shooter bias paradigm. *Journal of Experimental Social Psychology*, *44*(5), 1409-1413.

Warm, J.S (Ed.). (1984). *Sustained attention in human performance.* Chichester, UK: Wiley.

Whyte, J., Grieb-Neff, P., Gantz, C., & Polansky, M. (2006). Measuring sustained attention after traumatic brain injury: Differences in key findings from the sustained attention to response task (SART). *Neuropsychologia*,*44*(10), 2007-2014.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159-177.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 449-455.

Wilson, K. M., Head, J., de Joux, N. R., Finkbeiner, K. M., & Helton, W. S. (2015). Friendly fire and the sustained attention to response task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(7), 1219-1234.

Wilson, K. M., Russell, P. N., & Helton, W. S. (2015). Spider stimuli improve response inhibition. *Consciousness and Cognition*, *33,* 406-413.

Zelenski, J. M., & Larsen, R. J. (2000). The distribution of basic emotions in everyday life: A state and trait perspective from experience sampling data. *Journal of Research in Personality*, *34*(2), 178-197.

# Appendix A

## Word Lists for Verbal Recall Task

Table A1

Word lists

| **Word List 1** | **Word List 2** | **Word List 3** |
|---|---|---|
| Ankle | Piston | Lemon |
| Saloon | Butcher | Hamlet |
| Icebox | Fiord | Shotgun |
| Slipper | Typhoon | Abode |
| Infant | Nectar | Poster |
| Mucous | Harness | Cigar |
| Pudding | Reptile | Painter |
| Hostage | Lobster | Steamer |
| Banner | Rattle | Sunset |
| Bullet | Bandit | Costume |
| Sulfur | Pepper | Bagpipe |
| Doorman | Morgue | Banker |
| Locker | Trumpet | Spinach |
| Piano | Singer | Hairpin |
| Sunburn | Blister | Beggar |
| Missile | Jelly | Skillet |
| Thicket | Salad | Invoice |
| Monarch | Settler | Robber |
| Cowhide | Sultan | Kettle |
| Leopard | Fabric | Glacier |

Appendix B

NASA-TLX questionnaire

**NASA Task Load Index Questionnaire**

This questionnaire is designed to measure subjective workload. Please draw a cross at the point on the line which best reflects your answer. Think about the task that you have just completed when filling out this questionnaire.

**Mental Demand** How mentally demanding was the task?

**0 10 20 30 40 50 60 70 80 90 100**

**Very Low** **Very High**

**Physical Demand** How physically demanding was the task?

**0 10 20 30 40 50 60 70 80 90 100**

**Very Low** **Very High**

**Temporal Demand** How hurried or rushed was the pace of the task?

**0 10 20 30 40 50 60 70 80 90 100**

**Very Low** **Very High**

**Effort** How hard did you have to work to accomplish your level of performance?

**0 10 20 30 40 50 60 70 80 90 100**

**Very Low** **Very High**

# Appendix C

## Script for experiment

1. Read through information and consent form *5 mins*

The information sheet and consent form are on the desk. Once you have read through the information form, please sign the consent form if you still want to participate and fill out the demographic details on the back of the consent form. Let me know if you have any questions.

2. Explain verbal recall task and technology they are using *5 mins*

*Ask participants to take off watches and put phone on silent. Instruct them to not use phone during experiment.*

You have been assigned to the radio mikey/mobility app on the iPhone/no technology condition. Throughout the actual experiment you will hear words through the headphones. For each task, these words will either be nonsense words, or real words. If they are going to be real words, you will be instructed at the beginning of the task that you need to pay attention to them, and that every time you hear a word through the headphones, you need to radio/say aloud an associated word back, using the technology you have been given. The associated word is whatever word pops into your head first, and must be a different word to the one you heard. For example you might hear 'apple' through the headphones, and then you might choose to radio back 'pear' or 'orange'. At the end of the task you will be instructed to write down all the words you can remember, including the words you heard through the headphones and the words you said/radioed back. You do not need to remember the words in pairs, just write down what words you can remember. You may be doing the verbal recall task alongside another task, and you will be instructed if that is the case. I will now show you how to use the technology.

*Instructions for Mobility App*

You will be holding the ‘Taser’ in your dominant hand, so you will have to hold the phone in your other hand. Swipe into the phone and go to the app here (*point to where the app is).* I am “UC2”, so to radio me, tap on “UC2” and then hold down the “push to talk button” for the duration of your message. We stay paired for 30 seconds, so if you do not radio me within 30 seconds, it will time out. This means that you have to tap on “UC2” before radioing through again. If we are already paired, just push the “push to talk” button to radio through. *Check it is working.*

*Instructions for Radio Mikey*

You will be holding the ‘Taser’ in your dominant hand, so you will have to use the radio with your other hand. *Show how to clip it on.* You push the button on the side for the duration of your message. *Check it is working.*

3. Practice audio trial *4 mins*

Please put the headphones on. *Test the headphones are working and put to volume 65.* You will now do a practice audio trial to familiarize you with the task. Please radio/say an associated word back when you hear a word through the headphones. The start of all tasks in this experiment will be signalled with a high pitched tone followed by three low pitched tones, and the end of the task signalled by one high pitched tone. This audio track for the practice audio trial will also demonstrate these sounds to you. Let me know when you are ready to begin. *Begin practice trial. If words do not come radioed through to me, give the feedback at the end. Also give feedback if they do not follow instructions properly.*

4. General information *3 mins*

At the end of each task, you will be instructed to complete a questionnaire. The questionnaire is on the desk. Start with the one on top, and once you have completed it, fold it over. *Point to questionnaire and word recall sheet.* If you are instructed to recall as many words as possible, please do this until I tell you time is up, and do not move onto another task until instructed. I will be sitting over there for the experiment *(point to chair behind screens)* and I will give you instructions every now and then. Please keep all questions until the end of the experiment, unless you are unsure of what you are meant to do. You will also be instructed to take a short break after each task, and you need to remain seated in the chair for this.

*Now turn on the vision space screen and dim the lights.* We will begin the experiment. This is your 'Taser' gun *(turn flystick on and hand them flystick*). When it asks you to respond, pull the trigger at the front of the 'Taser' *(demonstrate).* Please put this hat on.

You will do two practice sustained attention trials before beginning the actual experiment. Please stand on the white cross on the floor.

5. Practice sustained attention trials

*Do not provide feedback on their performance in terms of omission or commission errors. Ensure all participants follow instructions of pointing to every stimuli- if they do not do this, then instruct them at the end of the practice trials* "point at every person, whether you are going to shoot them or not".

# Appendix D

## Descriptive statistics for variables

Table D1.

Mean correct response time (seconds) for each task across technology groups. Standard error in brackets

| | Target rich single | Target sparse single | Target rich dual | Target sparse dual |
|---|---|---|---|---|
| Control | 0.987 (.029) | 1.003 (.025) | 1.036 (.031) | 1.061 (.024) |
| Mobility App | 1.005 (.041) | 1.022 (.033) | 1.035 (.035) | 1.038 (.030) |
| Radio | 0.937 (.028) | 1.011 (.031) | 0.978 (.029) | 1.016 (.021) |

Table D2.

Mean percentage of commission errors for each task across technology groups. Standard error in brackets

| | Target rich single | Target sparse single | Target rich dual | Target sparse dual |
|---|---|---|---|---|
| Control | 6.60 (1.51) | 0.25 (.09) | 11.00 (1.60) | 0.58 (.15) |
| Mobility App | 7.00 (1.33) | 0.20 (.06) | 9.00 (1.50) | 0.70 (.13) |
| Radio | 13.60 (3.48) | .53 (.19) | 13.80 (2.08) | 1.00 (.24) |

Table D3.

Mean percentage of omission errors for each task across technology groups. Standard error in brackets

| | Target rich single | Target sparse single | Target rich dual | Target sparse dual |
|---|---|---|---|---|
| Control | 0.65 (.17) | 2.00 (.68) | 1.63 (.60) | 3.00 (.76) |
| Mobility App | 0.65 (.29) | 2.20 (.79) | 3.83 (1.03) | 5.60 (1.63) |
| Radio | 0.88 (.26) | 1.60 (.89) | 2.05 (.50) | 6.00 (1.18) |

Table D4.

Mean number of words recalled (out of 40) for each task across technology groups. Standard error in brackets

| | Verbal recall task | Target rich dual | Target sparse dual |
|---|---|---|---|
| Control | 20.55 (1.31) | 14.80 (1.04) | 16.50 (.98) |
| Mobility App | 14.10 (.92) | 14.70 (.85) | 14.05 (.79) |
| Radio | 17.32 (1.19) | 12.84 (.63) | 13.37 (.98) |

Table D5.

Mean mental demand ratings (out of 100) for each task across technology groups. Standard error in brackets

| | Control | Mobility App | Radio |
|---|---|---|---|
| Target rich single | 44.68 (4.93) | 45.08 (5.82) | 51.45 (4.87) |
| Target sparse single | 46.55 (4.78) | 40.88 (5.86) | 47.10 (5.28) |
| Verbal recall task | 62.95 (3.47) | 46.80 (4.98) | 57.85 (6.04) |
| Target rich dual | 72.45 (3.41) | 70.30 (4.59) | 76.20 (3.74) |
| Target sparse dual | 73.03 (3.13) | 70.80 (4.17) | 69.88 (3.92) |

Table D6.

Mean physical demand ratings (out of 100) for each task across technology groups. Standard error in brackets

| | Control | Mobility App | Radio |
|---|---|---|---|
| Target rich single | 33.34 (5.05) | 40.45 (5.37) | 36.23 (4.88) |
| Target sparse single | 33.76 (5.08) | 39.18 (5.54) | 33.85 (5.38) |
| Verbal recall task | 7.90 (1.80) | 14.75 (3.21) | 17.48 (3.55) |
| Target rich dual | 41.84 (6.13) | 47.18 (6.00) | 48.00 (4.93) |
| Target sparse dual | 35.71 (4.64) | 46.35 (6.66) | 40.38 (4.60) |

Table D7.

Mean temporal demand ratings (out of 100) for each task across technology groups. Standard error in brackets

| | Control | Mobility App | Radio |
|---|---|---|---|
| Target rich single | 44.05 (5.05) | 53.23 (4.76) | 48.75 (3.96) |
| Target sparse single | 43.80 (4.40) | 47.38 (6.02) | 42.53 (4.52) |
| Verbal recall task | 22.50 (4.57) | 27.25 (4.47) | 28.40 (5.40) |
| Target rich dual | 59.85 (4.69) | 70.08 (4.33) | 63.30 (3.60) |
| Target sparse dual | 54.90 (3.64) | 63.25 (5.33) | 62.13 (4.10) |

Table D8.

Mean effort ratings (out of 100) for each task across technology groups. Standard error in brackets

| | Control | Mobility App | Radio |
|---|---|---|---|
| Target rich single | 53.43 (4.73) | 53.15 (4.98) | 52.00 (4.27) |
| Target sparse single | 52.48 (4.27) | 48.18 (5.56) | 46.05 (5.07) |
| Verbal recall task | 54.70 (4.11) | 51.23 (5.79) | 53.70 (5.55) |
| Target rich dual | 72.13 (3.96) | 73.13 (3.71) | 74.00 (3.50) |
| Target sparse dual | 69.50 (3.11) | 72.85 (4.35) | 68.90 (3.20) |

# Appendix E

## ANOVA's for position from previous target

A 3 (technology group) X 5 (positions from previous target) repeated measures ANOVA was conducted for each task, with group as a between subjects variable.

There was a significant linear trend for position from previous target in the target rich single condition, $F(1, 56) = 363.13$, $p < .001$, $\eta_p^2 = .87$. There was a significant linear trend for position from previous target in the target sparse single condition, $F(1, 50) = 109.98$, $p < .001$, $\eta_p^2 = .69$. There was a significant linear trend for position from previous target in the target rich dual condition, $F(1, 56) = 301.77$, $p < .001$, $\eta_p^2 =.84$. There was a significant linear trend for position from previous target in the target sparse dual condition, $F(1, 46) = 82.24$, $p < .001$, $\eta_p^2 =64$. As seen in Figure 7, response time increases as distance from previous target increases.

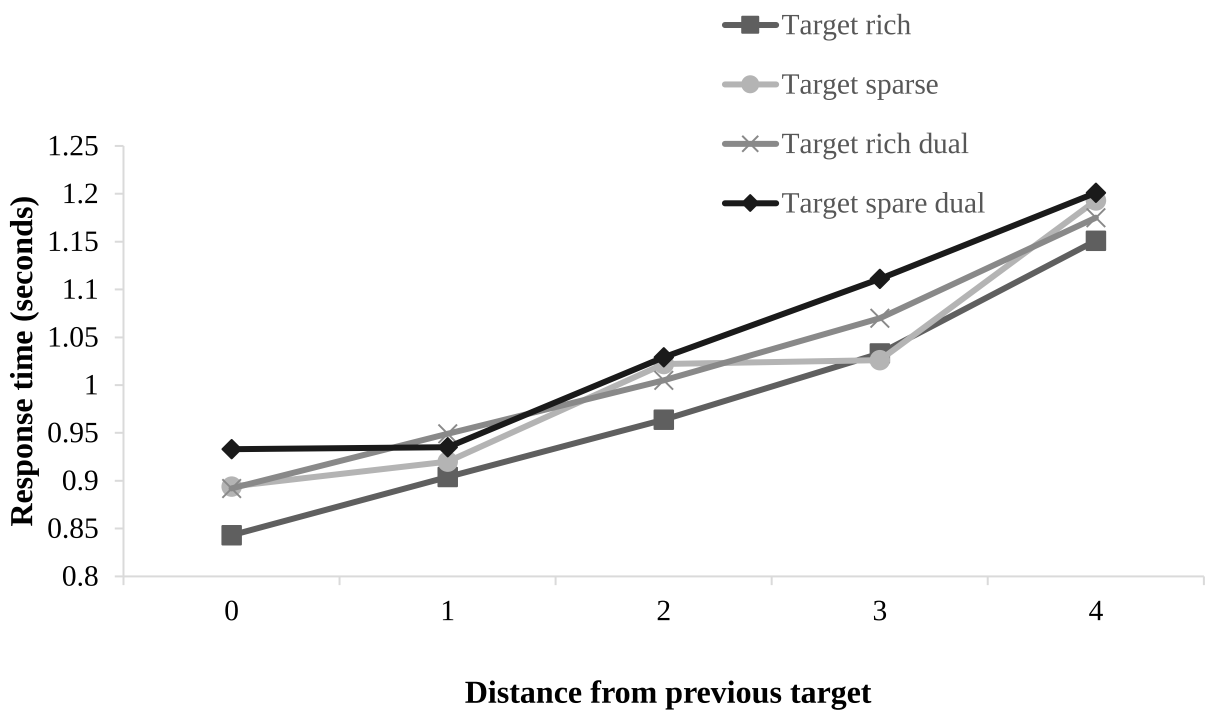


*Figure 7.* Linear trends for each task, showing that response time increases as position from previous target increases.